Cuk-H05899-16-P8576

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

JUNE 1975

VOL. 45, PART 2

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

Edited by
N. J. ENTWISTLE

Assistant Editors :
S. N. BENNETT
G. BROWN

Sub-Editor :
S. M. CUNNINGHAM

Advised by
L. B. BIRCH
R. A. C. OLIVER
E. A. PEEL
P. E. VERNON

Vol. 45, Part 2

The British Journal of Educational Psychology is issued by the British Psychological Society and the Association of Teachers in Colleges and Departments of Education, once a term, in February, June and November. The subscription is £4 a year post free, payable in advance, or £1·50 each number.

Subscriptions and Orders should be sent to The Manager, B.J.E.P. Department, Scottish Academic Press, 25 Perth Street, Edinburgh, EH3 5DW. Members of the British Psychological Society and members of the A.T.C.D.E. receive the *Journal* on special terms through the Societies.

Papers for publication should be sent to Professor N. J. Entwistle, Department of Educational Research, University of Lancaster, Lancaster. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for the *British Journal of Educational Psychology* should be sent to Professor N. J. Entwistle at the above address. *Advertisements* for this *Journal* should be sent to Scottish Academic Press.

Published by
SCOTTISH ACADEMIC PRESS, 25 PERTH STREET, EDINBURGH, EH3 5DW,
for the British Journal of Educational Psychology, Ltd.

Registered Offices : 8-24 Sampson Road North, Birmingham, B11 1BD. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of
The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

© 1975. The British Journal of Educational Psychology.

PRINTED BY BIRMINGHAM PRINTERS, LTD., 8-24 SAMPSON ROAD NORTH, BIRMINGHAM.

.5r

Br. J. educ. Psychol., 45, 111-121, 1975

370.1505
B777

EDUCATIONAL PSYCHOLOGY : ITS SCOPE AND LIMITS

By D. KALLOS
(*Institute of Education, University of Lund, Sweden*)
AND U. P. LUNDGREN
(*Institute of Education, University of Göteborg, Sweden*)

P8576

SUMMARY. An examination of current textbooks of educational psychology and educational research suggests that the scientific paradigm has been accepted too readily. This article explores the way in which the 'territory' of educational research has been defined and limited, by assumptions about scientific methods of enquiry and objective forms of measurement. It is argued that the reductionism found in current psychological explanations of the teaching-learning process has prevented fuller, and more realistic, descriptions in which the whole social context of teaching is considered, in terms of rules which govern what happens in classrooms.

ANY scientific discipline might on an epistemological level be described in terms of its territory (cf. Törnebohm, 1971a, 1971b). This article is concerned with educational psychology as a discipline and, therefore, with its territory. The discussion is based upon three assumptions. Firstly, we assume that any discipline is always concerned with some part of reality, and the delineation of that part defines its territory. Secondly, we assume that the production of knowledge within a given discipline aims at the creation of an increasingly refined or elaborated map of its territory. These two assumptions do not legitimate any specific epistemology. They represent one way of looking at research, and merely serve as brief statements of what any epistemology and scientific methodology stands for when applied to a certain discipline. Thirdly, we assume that educational psychology is a part of education. Our discussion will accordingly confront current production of knowledge within educational psychology as it affects the science of education.

(16)

The views on research in this article may be regarded as an attempt to apply an epistemological perspective developed by Törnebohm (1971a, 1971b). He regards research as "... a sequence of transformations of complexes composed by knowledge, problems and instruments (the term instrument refers to hardware tools used in laboratories and to software tools such as mathematical and statistical techniques). It will be assumed that research is concerned with a part of the real world. Knowledge may then be described as an authorised map over a territory." (Törnebohm, 1971a, p. 2.)

The ideas outlined above suggest that any scientific inquiry starts with some knowledge of the territory in question. That knowledge may at the onset be vague, and the boundaries of the territory may seem blurred. In this article we address ourselves specifically to the problems of educational psychology as a discipline. In terms of the frame of reference indicated above, the general aim of our discussion can be expressed in a question : what governs the production of knowledge in educational psychology ? In order to answer this broad question two more specific questions have to be asked :

(1) Which scientific ideals and perspectives govern the research efforts in educational psychology, i.e., which ideas on, for example, research methodology are prevailing within the field ?

(2) Which assumptions and perceptions of the territory govern the efforts to accumulate knowledge within educational psychology, that is, which questions are regarded as legitimate within the discipline ?

The very existence of the discipline of educational psychology answers the above questions to some extent. These answers may be found through an analysis of existing textbooks and research reports to which the label 'educational psychology' is affixed. Our aim is to articulate the answers already given, and to discuss their various consequences.

It stands to reason that the answers given are dependent on certain ideas, and even a scanty knowledge of the discipline of educational psychology suggests that there are several ways of defining its territory and its research strategies. This means that different answers to the questions posed may be conceived. We have, however, deliberately chosen to use, as a point of departure, currently accepted ideas of the field, as they are expressed in major sources from the United Kingdom and the United States. The production of knowledge within educational psychology is, as already mentioned, dependent on, among other things, scientific ideals and basic assumptions concerning the territory. The role of educational psychology within pedagogics as such, and the research efforts within educational psychology, is then dependent on the prevailing views on such issues. It follows that educational psychology in the United Kingdom and in the United States is different from educational psychology in, for example, the USSR (Simon and Simon, 1963 ; Mitter, 1974 ; Petrovsky, 1974). We will, however, try to answer the questions posed within the obviously accepted paradigm of educational psychology, so that our point of departure is within the frame of reference constituted by the "... recognised scientific achievements that for a time provide model problems and solutions to a community of practitioners." (Kuhn, 1970, p. viii.)

In order to provide some concrete examples of the consequences of the currently prevailing views we try to examine somewhat more closely the knowledge production within educational psychology as applied to problems concerning the teaching process. The final part of this article will thus try to answer the general question presented above as it is reflected within the area of relations between the research and the practice of teaching (Alkin, 1973).

What are the Scientific Ideals or Perspectives Governing Research in Educational Psychology ?

It is quite clear that research in educational psychology usually is regarded as a major part of educational research as such. The commonly accepted view further states that scientific research is governed by a set of rules and principles shared by all scientists, irrespective of the chosen discipline. Thus, all major textbooks on research in education tacitly accept a number of scientific ideals as common. These are presented without serious discussion or a shadow of a doubt (see, for example, Kerlinger, 1973 ; Travers, 1970). Common points of view are that research is an objective enterprise, and objectivity is defined by reference to accepted research methods. Furthermore, it is argued that there exists an objective observational language separated from the theoretical language. Translation between the two languages can be performed through the use of a set of translation rules. Through descriptions—correlational or experimental (Cronbach, 1957)—the researcher aims to establish laws or lawful relationships. The scientific ideals also include rules for the falsification (or verification) of hypotheses. These rules are essentially based upon the idea that the phenomena under study may be expressed quantitatively. The examples given here could be extended and elaborated. As our aim is not to discuss philosophy of science as such, it may suffice to state that ideals, of the kind mentioned here, are presented in standard textbooks without real attempts to justify them. Recently Suppes (1973) discussed "Facts and Fantasies of Education." He labelled as 'second-hand fantasies' beliefs about "... the efficiency or lack of it of a certain methodology, which is unsupported by evidence or systematic argument" (Suppes, 1973, p. 14). Although he generally accepted the support for experimentation in educational research advocated by Campbell and Stanley (1963) he was forced to admit that the authors, at least seemingly, were unable to produce real arguments in favour of their views. Interestingly enough, the discussion by Suppes (1973) is also void of any real arguments in favour of experimentation or quasi-experimentation in *educational* research. Other writers have accepted a narrow logical empirical definition of 'science' and 'scientific,' and sadly concluded that educational researchers have to remain unscientific (see, for example, Ebel, 1967). These examples imply that the standard notions on 'science' and 'scientific methodology' are far from self-evident and the consequences of the accepted ideals have wide implications.

Recently, several signs of criticism against the predominant views have become apparent (Thomas, 1972). The Anglo-American researchers are finally catching up with the debate on problems of pedagogics as a science already expressed in, for example, Germany (Klafki, 1971 ; Ulich, 1972). Ironically enough, logical empiricism had to fight for recognition within pedagogical research in post-war West Germany (Brezinka, 1959). Criticism about ideas of "Pädagogik als Geisteswissenschaft" as well as towards the ideals of logical empiricism has, of course, also been voiced by adherents of a critical theory approach and by Marxist theorists, such as Bourdieu and Passeron (1970), Huisken (1972), Suchodolsky (1972). The sources of reference indicated could imply that dissenting voices cannot be found in, say, the United Kingdom. This is, of course, not the case. But a vivid debate regarding the scientific ideals of educational research in general, and educational psychological research in particular, is certainly not very obvious. Books like that by Simon (1971) are an exception rather than a rule.

The dominating ideals briefly outlined above and regarded as governing research in educational psychology may, of course, be compared with contemporary research in the hard sciences which apparently serve as somewhat diffuse but ideal models for authors of textbooks in educational research methodology. The theoretical biologist or physicist (Waddington, 1968) would most certainly feel somewhat alienated and frustrated in trying to adhere to the scientific ideals originally attributed to him and his colleagues. This would be true even, to use Scriven's (1972) words, for "... that most backward and traditional field, medicine ..." (p. 116). At the concrete level the influence of a particular approach to science might be described by elaborating the analogy of mapping out a territory. The scientific perspective defines the function of the map, and thus answers the question of the use of the map.

The function of knowledge in educational psychology seems to be at least twofold. Firstly, it tries to help the researcher, the student of education, and the teacher to obtain a picture of the landscape. Through careful studies of details the student will be trained to observe what is important and to sort out what is not. Secondly, it provides advice concerning how to move from one point to another on the map, and it also includes a comprehensive guide as to what places are worthwhile visiting.

The first type of map may be illustrated by statements commonly found in introductions to textbooks in educational psychology. Ausubel (1968), for instance states : " The basic premise underlying this book is that educational psychology is primarily concerned with the nature, conditions, outcomes, and evaluation of classroom learning " (p. vii), while Child (1973) presents the aim of his textbook in the following way : " The central aim of the book is to introduce teachers to elementary ideas in psychology which have some relevance for their work with young children" (p. v). Finally, Stephens and Evans (1973) state : " This book deals with the psychology of the important processes at work within the classroom. It attempts to use psychology to help the teacher in understanding the developing student, in directing academic achievement, in promoting the development of character, and in aiding the attainment of personal adjustment " (p. v).

The second idea concerns all kinds of prescriptive advice to be found in textbooks like those quoted above. The ideals outlined here have an influence in another way, too, namely, on the nature of the map. They tend to govern researchers on questions of content, that is, they have a bearing on questions concerning what territory should be claimed and mapped out. Furthermore, through legitimising certain instruments there exists a relation between data-gathering procedures and the resulting map. The stress is on quantification and on a high degree of generalisability. If, for example, ' mountains ' are to be included, then instruments must be constructed to measure their height, and when such instruments have been developed, there is a tendency always to include mountains in the maps, regardless of the purpose of the map. A tendency to perceive mountains and to look for them will also be demonstrated, thereby perhaps disregarding other aspects of the territory. To some extent the lack of awareness of underlying ideals and perspectives that is mirrored in textbooks on research methodology in pedagogics can be explained in terms of long chains of transformations between problems and methods, leading to a reification of the problems through perceiving them as purely methodological. Marzahn (1971, p. 32) illustrates this point nicely by stating : " . . . one rationalises knowledge in detail (Method)—and is defenceless against the historical irrationality of the aims and purposes to which it is subservient." (Translation). This leads to our second question.

What Assumptions and Which Perceptions of the Territory Govern the Attempts to Gather Knowledge in Educational Psychology ?

The basic assumption underlying the definition of the territory of educational psychology may be stated, in its simplest form, in the following way. Education deals with the implementation of knowledge, attitudes and skills in growing individuals through the processes of teaching and upbringing. The selection of content is *not* a scientific question for the researcher within the field of educational psychology. The only scientific aspect of the problem of selection of contents is the possibility for a given individual or group of individuals to learn the tasks in question at, for example, the age level suggested in the curriculum guide. In practice, it is considered ' scientific ' to study relations between individual parameters and parameters referring to content thus aiming to establish such things as ' prerequisites for learning ' (Gagné, 1970, 1972).

Educational psychology is then defined as the application of *selected portions* of psychological knowledge to the narrowly perceived problems of education. The task of educational psychology is to present solutions in the form of lawful relationships between three sets of parameters and their eventual interaction : (1) goal (or content) parameters ; (2) individual parameters ; and (3) situational parameters. This task is described in a more formalised way by

Walberg (1971). The model proposed by him is specifically applied to questions of optimising and individualising learning, but it is sufficiently general to serve as an illustration. His general instructional model is multiplicative and, in brief, it states that learning something specific (L_h) is the function (b) of aptitude (A_i), environment (E_j), instruction (I_k), and the interaction between these. This is expressed in the following equation :

$$\text{“}L_k=b_1(A_i)+b_2(E_j)+b_3(I_k)+b_4(A_iE_j)+$$

$$+b_5(A_iI_k)+b_6(E_jI_k)+b_7(A_iE_jI_k).\text{”}$$

It certainly looks very impressive and scientific, but what does it really stand for ?

Another example is provided by the review of educational research presented by Kerlinger (1969). He boldly states that educational research is psychological research applied to educational problems—because an overwhelming majority of its variables are psychological or socio-psychological. His conclusion is correct if the criterion is limited to the variables currently used. However, he is incorrect in assuming that this is necessarily the proper state of affairs, as, to use our analogy, he is defining the territory solely from the existing maps !

A third example is provided by a recent text on the relations between psychology and education (Hunt and Sullivan, 1974). The authors recognise the fact that psychological research has a none too glorious history when it comes to applications to educational problems. But Hunt and Sullivan see education as being mainly, if not solely, a practical endeavour. Educational research, *per se*, is accordingly not even mentioned in the book. Thus, it is implicitly understood that psychological research directed at educational problems is a prime source for educational practice, and perhaps even that educational psychological research *is* educational research. It might then follow that no such thing as a scientific discipline of education exists in its own right.

The main assumption guiding educational psychologists seems to be that educational problems may be reduced to psychological ones. Several concrete examples of such ‘ reductionism ’ may be mentioned. At a more concrete level, the treatment of motivational problems in education may be offered as an example. Rather than discussing that problem in terms of the system developed by Maslow (1954), or in terms of a reinforcement model (Bijou, 1970) which is standard today, we would suggest a more direct approach. Such an approach should focus on the relations of *what* is being taught, not to drives or motivational constructs, but to the student's present knowledge and life situation, to his possibilities of using his acquired knowledge in a manner meaningful to him, and to how new knowledge may be used to change his own situation. Esland (1971) provides a stimulating discussion of how different conceptualisations of man and knowledge have different consequences for, among other things, views on classroom motivation. His discussion may be considered within the perspective outlined here.

Nevertheless, educational concepts are still to a considerable extent psychological, and the attempts to develop an ‘ educational language ’ of a ‘ scientific nature ’ are still rather feeble (see, for example, Scheffler, 1960 ; MacMillan and Nelson, 1968 ; Hirst and Peters, 1970). We have argued elsewhere (Kallós, 1973a, 1973b ; Lundgren, 1973a) that a conceptual development in the area of teaching is one product of meaningful scientific research, but not a necessary prerequisite for such research (Russell, 1948). It is an illusion that education (or pedagogics) doesn't have any territory of its own. When we start to map out education as a series of phenomena of distinctive nature (Gowin, 1972) we will gain knowledge that can be used in order to translate psychology into educational psychology. Basically, the same notion was stated earlier by Smith (Smith and Meux, 1962 ; Smith, 1967) when he discussed the relations between learning and teaching. Today, educational psychology as a scientific discipline works with a territory of very limited relevance for education (or pedagogics) both as science and as practice.

Educational Psychology and the Problems of Teaching.

Teaching (or instruction) is considered to be a key concept in education. In many instances educational psychology is simply defined as the application of psychological findings and ideas to that activity. In its turn teaching is regarded as an activity concerned with classroom learning. Educational psychologists tend to regard the links between learning and teaching as strong. Reductionalistically viewed, teaching simply means the arrangement of external conditions for learning (Gagné, 1970) derived from psychological knowledge of the learning process. The relations between external processes (teaching or instruction) and internal processes (learning) are thus conceived as direct and straightforward. Of course, this conceptualisation of teaching and its relation to learning has been the subject of serious debate, excellently summarised by Westbury (1971).

The problems involved in relating the external processes of teaching to internal psychological processes may be illustrated by some examples. Educational technology in its orthodox form tries to reconstruct teaching, so that this activity will conform to principles of learning. It then legitimises a far from non-controversial separation between means and ends in teaching, and a very narrow view of the social and economico-political objectives of teaching within a society. It could, for example, be argued that the main functions of schooling and teaching are to provide society with a selection and qualification system as a part of the social reproduction of that society. From such a perspective the conclusions and recommendations of educational technologists must imply a severe reductionism.

Research on interactions between aptitudes and treatments (Cronbach, 1967 ; Cronbach and Snow, 1969 ; Bracht, 1970 ; Salomon, 1972 ; Berliner and Cahen, 1973) might be used as another example. Here teaching is conceived as treatment and is presumably to be designed by taking into account relations between aptitude patterns and learning. Well within the limits of educational psychological research traditions, the aptitude-treatment interaction studies can be severely criticised if a pedagogical perspective is chosen (Kallós, 1975). The concept and the ideas of instruction used by researchers in the area of aptitude-treatment interaction is, to mention just one example, very far away from what is actually possible to implement or even desirable to introduce in our school systems. Many of the attempts of educational psychologists to describe and analyse the teaching process *per se* have, however, only looked upon this process as an intervening variable between antecedents and outcomes. The futility of such an approach for education was recognised by Shulman (1970). His suggestions for a necessary re-orientation, however, did not build upon close analysis of the nature of educational phenomena. Still other researchers have focused upon the relations between process and outcomes, disregarding important questions of *how* and *why* a certain teaching process was implemented (Rosenshine, 1971), or simply explained such issues in terms of teacher traits.

Even when actual process research has been carried out, the categorisation of teaching events has almost without exception relied on psychological or socio-psychological classification systems. The most noteworthy exceptions to this rule are the systems developed by Smith and Meux (1962) and by Bellack *et al.* (1966), where a pedagogical analysis of teaching is taken as a starting-point. The scientific ideals of educational psychologists studying teaching also made them rely heavily upon the quantifiable aspects of teaching, disregarding the important illuminative descriptions in qualitative terms provided by Jackson (1968) and Smith and Geoffrey (1968) in their analyses. The reduction by educational psychologists has already been done in the choice of a category system, and thus this type of research nicely illustrates the effects on the resulting maps of a chosen perspective of the territory in question.

A final example is provided by the survey of teaching during fifty years presented by Hoetker and Ahlbrand (1969). The demonstrated stability of teaching patterns indicates that all the efforts of educational psychologists and curriculum thinkers to change the pattern of teaching were more or less in vain. The rules that shape the teaching survived irrespective of normative advice. The territory remained intact, irrespective of the maps. But if we start to reflect on findings like those presented above, we must begin to doubt the claim by educational psychologists that teaching as a phenomenon can be reduced to a psychological process.

Pedagogically speaking, teaching has to be looked upon as an integral part of the educational system. The starting point is thus the system, and not the psychological processes within the individual learner or the teacher. From a theoretical point of view this will lead to a set of concepts not derivable from psychology, and to a refutation of logical empiricism as the *sole* scientific basis for educational research. The examples mentioned indicate that the production of knowledge within educational psychology is not governed by ideas about teaching, but rather on the assumed possibility to *deduce* workable hypotheses and suggestions about the external process of teaching from a body of knowledge in psychology, and from a set of rules on how to conduct acceptable scientific research.

It is customary to attack any criticism if it can be said to be only destructive i.e. if the critic does not offer constructive solutions. But our primary aim is not to provide new model solutions to old problems, but rather to advocate that educational psychologists start to reflect more consciously on the basic assumptions governing their research efforts. We have tried to make some of these assumptions, ideals, and perspectives explicit as a starting point for that necessary reflection. We have also tried to give examples of the consequences of these assumptions and ideals. Several paths seem to be open for a redirection. While the aim of this article was not to provide our own solutions to the dilemma, it is possible to outline some aspects of *one* possible framework for the study of teaching, by examining teaching as a rule governed activity. The rules are socially determined, and thus not formulated in terms of inferred psychological processes. If this is the case, we must try to answer questions regarding the nature of the rules, and how these rules are generated, what consequences they have, and under what circumstances they might be changed. Following ideas originally presented by Dahllöf, (1971, 1974) we have tried to develop a conceptual framework for the study of schooling and teaching, and to conduct empirical research along these lines (Kallós, 1973a, 1973b ; Lundgren, 1972, 1973a, 1973b, 1974 ; Björklund, Kallós and Larsson, 1974).

According to our views the teaching process is regarded as constrained by various externally imposed 'frames.' The frame factors define the space of options available to teachers and students. Frames may be decided upon at various levels within the educational system (Alkin, 1973 ; Dahllöf, 1974).

At the classroom level frame decisions concern the organisation, time, contents and persons involved in teaching. In their turn these decisions are constrained and directed by decisions at higher levels within the educational system.

If teaching in a given subject during a certain short time period is chosen as an example, the content frames may include decisions about textbooks to be used, and objectives to be pursued (in a given case perhaps even formulated in behavioural terms). Organisational frame decisions may concern the composition and grouping of the school class, and personal frame factors may refer to decisions regarding the selection of teachers and students, by, for example, the assignments of a particular teacher to a particular group of students. The actual scope of a decision at a given level within the educational system is determined by the rules (laws) of that particular system. Teachers are, or are not, permitted to decide by themselves upon the use of a certain textbook, to mention but one example.

The concept of ' frame ' has also been used by Bernstein (1971). To him this concept refers to " . . . the degree of control teacher and student possess over the selection, organisation and pacing of knowledge transmitted and received in the pedagogical relationship " (p. 50). In our writings we have used the concept of frame to cover other aspects of the constraints on the pedagogical relationship in the classroom, too. We quite agree, however, that content frames are of prime importance in the establishment of rules operating in the teaching process. The classification and framing of knowledge (contents), using Bernstein's (1971) terms, explicitly reflect the image of man and the very concept of knowledge applied within a particular educational system.

Using the approach of studying teaching as a socially determined process constrained by frames imposed, means that we leave the area of 'lawful relationships' in the positivist meaning of this concept. Instead we look upon how certain rules operating in the 'game' of teaching are formulated as a consequence of a given set of frame factors. The activities actually taking place in the classroom represent more or less conscious solutions to the problems perceived by teachers and students. Certain "moves" or sequences of "moves" are more or less plausible under a certain set of frames. It has, for example, been shown that the communication patterns, the teaching tempo (length of time devoted to a certain curriculum unit) within the class, may be steered by a certain group of pupils. Under the *particular* frame conditions investigated, this steering group (or ' criterion group ' if we use the teacher's perspective) was formed by the students between the 10th and the 25th percentile with regard to ability (Lundgren, 1972, 1974). The fact that the teachers studied used such a steering-group should be interpreted in terms of the frames imposed within the Swedish school system, and be regarded as a solution to the problems facing the teacher under those circumstances. A change in frame conditions would probably alter the position of the steering-group, or it would perhaps not exist at all. There already exist some data from studies at the university level to support such a conclusion (Björklund, Kallós and Larsson, 1974). The brief concrete example offered above could, if regarded at another level, serve as an example of the problems of educational planning. Were the teaching strategies actually observed in the classroom really expected or wanted by the planners at higher levels ?

In our work we have come to regard conventionally applied psychological variables, or traditional approaches derived from research traditions within educational psychology, as insufficient and rather trivial in the attempts to study the territory of education. In this article we have tried to state some of the reasons for our discontent. A proper discipline called educational psychology, springing from the well cultivated garden of psychology, must, in our view, include a fund of *educational knowledge*, as well as basic psychological theories, if it is to be successfully transplanted into the school-yard.

ACKNOWLEDGMENTS—The writing of this paper was supported by grants from the Department of Educational Research and Development of the Office of the Chancellor of the Swedish Universities, as part of a research project concerning the " Relative Merits of Various Teaching Methods in Higher Education," and by the Tercentenary Fund of the Bank of Sweden, as part of a research project concerning " Model Analysis of Pedagogical Processes." The former project is carried out at University of Lund and the latter at the University of Goteborg. The points of view, selection of facts, and opinions expressed in this article are those of the authors and do not necessarily coincide with the official positions of the funding agencies.

REFERENCES

ALKIN, M. C. (1973). Theoretical framework for the analysis of curriculum and instructional reform. *Int. Rev. Educ.*, 19, 195-207.

AUSUBEL, D. P. (1968). *Educational Psychology : A Cognitive View.* New York : Holt, Rinehart and Winston.

BELLACK, A. A., KLIEBARD, H. M., HYMAN, R. T., and SMITH, F. L. (1966). *The Language of the Classroom.* New York : Teachers' College Press.

BERLINER, D. C. and CAHEN, L. S. (1973). Trait-treatment interaction and learning. In KERLINGER, N. F. (Ed.), *Review of Research in Education, Vol.* 1, 58-94. Itasca, Ill. : F. E. Peacock.

BERNSTEIN, B. (1971). On the classification and framing of educational knowledge. In YOUNG, M. F. D. (Ed.), *Knowledge and Control*, 47-69. London : Collier-Macmillan.

BIJOU, S. W. (1970). What psychology has to offer education—now. *J. appl. Behav. Analysis.*, 3, 65-71.

BJÖRKLUND, S., KALLÓS, D., and LARSSON, L. (1974). Hardware gone soft. Teaching under strong frame conditions at an institute of technology. *Report from the Institute of Education, University of Lund, Sweden*, 48. Mimeo.

BOURDIEU, P., et PASSERON, J.-C. (1970). *La reproduction : Elément pour une théorie du système d'enseignement.* Paris : Editions de Minuit.

BRACHT, G. H. (1970). Experimental factors related to aptitude-treatment interactions. *Rev. educ. Res.*, 40, 627-646.

BREZINKA, W. (1959). Die Pädogogik und die erzieherische Wirklichkeit. *Zeitschrift für Pädagogik*, 5, 1-34.

CAMPBELL, D. T., and STANLEY, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In GAGE, N. L. (Ed.), *Handbook of Research on Teaching*, 171-246. Chicago : Rand McNally.

CHILD, D. (1973). *Psychology and the Teacher.* Old Woking, Surrey : Holt-Blond.

CRONBACH, L. J. (1957). The two disciplines of scientific psychology. *Am. Psychol.*, 12, 671-684.

CRONBACH, L. J. (1967). How can instruction be adapted to individual differences ? In GAGNÉ, R. M. (Ed.), *Learning and Individual Differences.* 23-29. Columbus, Ohio : C. E. Merrill.

CRONBACH, L. J., and SNOW, R. E. (1969). *Final Report : Individual Differences in Learning Ability as a Function of Instructional Variables.* School of Education, Stanford University, Standford, California. Mimeo.

DAHLLÖF, U. (1971). *Ability Grouping, Content Validity, and Curriculum Process Analysis.* New York : Teachers' College Press.

DAHLLÖF, U. (1974). Trends in process-related research at different problem levels in educational sciences.. *Scand. J. educ. Res.*, 19, 55-77.

EBEL, R. L. (1967). Some limitations of basic research in education. Reprinted in BROUDY, H. S., ENNIS, R. H. and KRIMMERMAN, L. I. (Eds.), *Philosophy of Educational Research*, 122-130. New York : Wiley, 1973.

ESLAND, G. M. (1971). Teaching and learning as the organisation of knowledge. In YOUNG, M.F.D. (Ed.), *Knowledge and Control*, 70-115. London : Collier-Macmillan.

GAGNÉ, R. M. (1970). *The Conditions of Learning.* (2nd edit.). New York : Holt, Rinehart and Winston.

GAGNÉ, R. M. (1972). Domains of learning. *Interchange*, 3 (1), 1-8.

GOWIN, D. B. (1972). Is educational research distinctive ? In THOMAS, L. G. (Ed.), *Philosophical Redirection of Educational Research*, 9-25. Chicago : University of Chicago Press.

HIRST, P. H., and PETERS, R. S. (1970). *The Logic of Education*. London : Routledge and Kegan Paul.

HOETKER, J and AHLBRAND, W. P. (1969). The persistence of recitation. *Am. educ. Res. J.*, 6, 145-168.

HUISKEN, F. (1972). *Zur Kritik bürgerlicher Didaktik und Bildungsökonomie*. München : List Verlag.

HUNT, D. E., and SULLIVAN, E. V. (1974). *Between Psychology and Education*. Hinsdale, Ill. : Dryden Press.

JACKSON, P. W. (1968). *Life in Classrooms*. New York : Holt, Rinehart and Winston.

KALLÓS, D. (1973a). On educational scientific research. *Report from the Institute of Education, University of Lund, Sweden*, 36. Mimeo.

KALLÓS, D. (1973b). Some comments on frame and control as educational concepts. In *Reports from the Institute for Educational Research, University of Jyväskylä, Finland*, 213/1973, 65-81. Mimeo.

KALLÓS, D. (1975). Comments on the importance of aptitude-treatment interaction research for pedagogics and for the theory and practice of teaching. In SCHWARZER, R., and STEINHAGEN, K. (Eds.), *Adaptiver Unterricht. Interaktionen zwischen Lehrmethoden und Lernermerkmalen*. München : Kösel Verlag.

KERLINGER, F. N. (1969). Research in education. In EBEL, R. L. (Ed.), *Encyclopaedia of Educational Research* (4th edit.), 1127-1144. New York : Macmillan.

KERLINGER, F. N. (1973). *Foundations of Behavioural Research* (2nd edit.). London : Holt, Rinehart and Winston.

KLAFKI, W. (1971). Erziehungswissenschaft als kritisch-konstruktive Theorie : Hermeneutik—Empirie—Idiologiekritik. *Zeitschrift für Pädogogik*, 17, 351-385.

KUHN, T. S. (1970). *The Structure of Scientific Revolutions* (2nd edit.) Chicago : University of Chicago Press.

LUNDGREN, U. P. (1972). *Frame Factors and the Teaching Process*. Göteborg : Almqvist and Wiksell.

LUNDGREN, U. P. (1973a). Educational theory. Like in a mirror image or by one's own will. In *Reports from the Institute for Educational Research, University of Jyväskylä, Finland*, 213/1973, 25-43. Mimeo.

LUNDGREN, U. P. (1973b). Pedagogical frames and the teaching process. A report from an empirical curriculum project. *Report from the Institute of Education, University of Göteborg, Sweden*, 25. Mimeo.

LUNDGREN, U. P. (1974). Pedagogical roles in the classroom. In EGGLESTON, J. (Ed.), *Contemporary Research in the Sociology of Education*, 200-213. London : Methuen.

MACMILLAN, C. J. B., and NELSON, T. W. (Eds.) (1968). *Concepts of Teaching*. Chicago : Rand McNally.

MARZAHN, C. (1971). Zur Bedeutung der revolutionären Pädagogen der Weimarer Republik für die Rekonstituierung einer marxistischen Erziehungstheorie und Praxis heute. *Erziehung und Klassenkampf*, 1 (1), 31-49.

MASLOW, A. H. (1954). *Motivation and Personality*. New York : Harper and Row.

MITTER, W. (Ed.) (1974). *Didaktische Probleme und Themen in der UdSSR*. Hannover : H. Schroedel Verlag.

PETROVSKY, A. V. (1974). Pedagogy and psychology. *Soviet Education*, 16 (5), 61-68.

ROSENSHINE, B. (1971). *Teaching Behaviours and Student Achievement*. Hove : National Foundation for Educational Research in England and Wales.

RUSSELL, B. (1948). *Human Knowledge. Its Scope and Limits*. New York : Simon and Schuster.

SALOMON, G. (1972). Heuristic models for the generation of aptitude-treatment interaction hypotheses. *Rev. educ. Res.*, 42, 327-343.

SCHEFFLER, I. (1960). *The Language of Education*. Springfield, Ill. : C. C. Thomas.

SCRIVEN, M. (1972). Objectivity and subjectivity in educational research. In THOMAS, L. G. (Ed.), *Philosophical Redirection of Educational Research*, 94-142. Chicago : University of Chicago Press.

SHULMAN, L. S. (1970). Reconstruction of educational research. *Rev. educ. Res.*, 40, 371-396.

SIMON, B. (1971). *Intelligence, Psychology and Education. A Marxist Critique*. London : Lawrence and Wishart.

SIMON, B. and SIMON, J. (Eds.) (1963). *Educational Psychology in the USSR*. London : Routledge and Kegan Paul.

SMITH, B. O. (1967). Toward a theory of teaching. In BELLACK, A. A. (Ed.), *Theory and Research in Teaching*, 1-11. New York : Teachers' College Press.

SMITH, B. O., and MEUX, M. O. (1962). *A Study of the Logic of Teaching* (Trial edit.). Urbana, Ill. : College of Education, University of Illinois. Mimeo.

SMITH, L. M., and GEOFFREY, W. (1968). *The Complexities of an Urban Classroom.* New York : Holt, Rinehart and Winston.

STEPHENS, J. M., and EVANS, E. D. (1973). *Development and Classroom Learning. An Introduction to Educational Psychology.* New York : Holt, Rinehart and Winston.

SUCHODOLSKI, B. (1972). *Einführung in die marxistische Erziehungstheorie.* Köln : Pahl-Rugenstein.

SUPPES, P. (1973). Facts and fantasies of education. In WITTROCK, M. C. (Ed.), *Changing Education.* 6-45. Englewood Cliffs, N.J. : Prentice Hall.

THOMAS, L. G. (Ed.) (1972). *Philosophical Redirection of Educational Research.* Chicago : University of Chicago Press.

TÖRNEBOHM, H. (1971a). *Reflections on Scientific Research.* Gothenburg : Institute for Theory of Science, University of Gothenburg, Sweden. Mimeo.

TÖRNEBOHM, H. (1971b). *Research as an Example of an Innovative System.* Gothenburg : Institute for Theory of Science, University of Gothenburg, Sweden. Mimeo.

TRAVERS, R. M. W. (1970). *An Introduction to Educational Research* (3rd edit.). New York : Macmillan.

ULICH, D. (1972). Probleme und Möglichkeiten erziehungswissenschaftlicher Theorienbieldung In ULICH, D. (Ed.), *Theorie und Methode der Erziehungswissenschaft*, 13-87. Weinheim und Basel : J. Beltz Verlag.

WADDINGTON, C. H. (1968). *Towards a Theoretical Biology.* 1. *Prologomena.* Edinburgh : University of Edinburgh Press.

WALBERG, H. J. (1971). Models for optimising and individualising school learning. *Interchange*, 2 (3), 15-27.

WESTBURY, I. (1971). Problems and prospects. In WESTBURY, I., and BELLACK, A. A. (Eds.), *Research into Classroom Processes. Recent Developments and Next Steps*, 227-252. New York : Teachers' College Press.

(*Manuscript received 18th November*, 1974)

Br. J. educ. Psychol., 45, 122-129, 1975

PUPIL BEHAVIOUR IN JUNIOR CLASSROOMS

By DEANNE BOYDELL

(*School of Education, University of Leicester*)

SUMMARY. The nature of children's behaviour in six informal junior school classrooms was examined using an observation schedule modelled on Medley's Personal Record of School Experiences. The classes were unstreamed, the children sat in small groups and there was a high incidence of private teacher talk to small groups and individual children. The observer focused on children one at a time using a time sampling procedure and recorded his observations by multiple coding. Children spent about a third of their time interacting with other children and rather more than half the lesson out of contact with everyone. Despite a low incidence of teacher contact children were predominantly involved in their work.

INTRODUCTION

CLASSROOM observation of young children was pioneered by Thomas *et al.* in 1929 (Medley and Mitzel, 1963). However, despite a long history of research interest and the current popularity of classroom observation, relatively few techniques focus on pupil behaviour as distinct from teacher behaviour and teacher-pupil interaction (Simon and Boyer, 1970). Instruments suitable for observing young children in informal classrooms are particularly rare. However, one such instrument has been developed for use with nursery school children in the USA. Called PROSE (Personal Record of School Experiences) (Medley *et al.*, 1973), it aims to " compile a record analogous to an ecological one but precoded . . . " (Medley, 1969).

PROSE is used by a 'live' observer who focuses on children one at a time in a prearranged arbitrary order and codes his behaviour across a number of relatively low-inference categories. Each child is observed five times, each occasion coinciding exactly with time signals which occur at 25-second intervals. On each occasion his behaviour is recorded on the ' Statement ' side of the record sheet. At the end of the 100-second period the observer codes the child's ' Context ' on the reverse side of the sheet, noting for instance the size of his group and the materials he has used. By collating the observations made on individual children it is possible to produce a highly detailed and comprehensive picture of pupil behaviour within any type of classroom, irrespective of the degree of organisational complexity, mobility or interaction.

The purpose of the exploratory study reported here was to investigate the nature of children's behaviour in unstreamed, informal, junior classrooms where the children sit in small groups and the incidence of private teacher talk to small groups and individual children is high. An observation schedule modelled on the ' Statement ' side of PROSE was used to examine the nature and frequency of children's classroom interactions and their activities when not interacting. Interest focused on the extent to which children's observable behaviour matched up to the popular image of informal classrooms. The findings supplemented the results of a study of teacher-pupil contact which was carried out simultaneously in the same classrooms using a specially designed teacher instrument (Boydell, 1974).

METHODS

Observation Schedule.

Six observers, all teachers on a full-time advanced diploma course, were trained to use the observation schedule. The observer focused on individual children in the class one at a time, in a predetermined arbitrary order. He coded each child's behaviour five times on an individual record sheet at the *precise* time of prerecorded time signals which were fed into his ear every 25 seconds from a portable cassette tape-recorder. After five signals the same procedure was repeated with the next child and so on. Working in this fashion the observer covered about 25 children in an hour. At the end of each visit the observer wrote a thumbnail sketch of the lesson, noting for instance whether there was a single task for all the children and whether apparatus was involved.

Each child's record sheet provides space for personal details, such as name and sex, and a column for each of the five time signals. In addition, it contains the 13 categories and their constituent items which are given in Tables 2, 3 and 4 (see later). The majority of the categories are similar or identical to those used on the ' Statement ' side of PROSE, although 8 is expanded and 3b, 7a, 9a and 9b are additions.

The target child's behaviour is coded by the method of multiple coding. The total number of codings made at a signal depends on the child's situation. Categories 8, 9a and 9b are always coded. In addition categories 1, 2, 3a and 3b are coded if the child is interacting with an adult, and categories 4, 5, 6, 7a, 7b and 7c are coded if the target is interacting with another pupil. The observer codes by ticking the one appropriate item within each relevant category. Consider, as an example, the following sequence : at the first signal the target child, a boy, is working by himself at his own table whilst the teacher helps another child elsewhere in the room. Moving to the first column the observer ticks COOP TK in 8, P IN in 9a and T ELSE (I) in 9b (see Table 4). At the second signal the teacher is out of the room and the target is initiating a work conversation with a boy at his own table who has the same task as himself. Working in the second column the observer ticks INIT in 4, COOP in 5, VRB in 6, S TK in 7a, SS in 7b, OWN GP in 7c, COOP TK in 8, P IN in 9a and T OUT RM in 9b (see Tables 3 and 4). The observer continues in like manner for signals three, four and five.

At the end of the training and practice period the reliability of the observers' coding was assessed using a 17-minute videotape test consisting of four short extracts of a child's classroom behaviour with time signals superimposed every 25 seconds. The average percentage agreement with the investigator on the child's situation was 94 per cent. On specific categories it was found that the agreements with the investigator were 67 per cent, 75 per cent and 76 per cent for categories 1, 3a and 3b, respectively ; 61 per cent, 68 per cent, 88 per cent, 82 per cent, 77 per cent and 88 per cent for categories 4, 5, 6, 7a, 7b and 7c, respectively, and 60 per cent for category 8. Categories 2, 9a and 9b were not coded in this videotape test because all the pupil-adult interaction was with the teacher, the target child never left his seat and the teacher's location and activity were not visible, except when interacting with the target's group. Six weeks later, at the end of the field-work, the observers were again tested with the same videotape with very similar results. It is relevant to note that the observers reported great difficulties in coding child behaviour from videotape : without exception they found ' live ' observation in classrooms far easier, the major reason being the ambient noise associated with the sound recording.

Data Collection.

The teacher sample consisted of six volunteer junior school teachers working in six different unstreamed city schools (see Boydell, 1974). All the classes were engaged in mathematics for the duration of each observational visit and all the rooms were arranged informally with the majority of children sitting in small groups. Four of the teachers usually provided a variety of activities (as opposed to one set task) and four of them frequently included some practical work. All relied heavily on individual and group attention in preference to class teaching.

Each of the six teachers was visited once a week on a Wednesday between the mid-morning break and lunchtime by one of the six observers in turn. This rotation of observers meant that every classroom was seen once by every observer and that 36 mathematics lessons of approximately an hour's duration were available for analysis. Each observer was accompanied on his visit by a second observer who focused on the teacher using a specially designed observation instrument described elsewhere (Boydell, 1974).

Data Processing.

The data from each lesson were processed separately using the SPSS computer programs (Nie, Bent and Hull, 1970). The incidence of each situation and category item were listed and a number of cross-tabulations were run. The results from the six separate lessons observed in each classroom were then collated to give a classroom profile. In the interests of consistency, classroom code letters in this paper correspond to the teacher code letters in the previous paper (Boydell, 1974) : classroom A was taught by teacher A and so on.

RESULTS

Type of Situation.

The frequencies of the three situations in which children were found were expressed as percentages of the number of recorded events for each lesson. The percentages for the six lessons observed in each classroom were averaged. The results are shown in Table 1. A two-way analysis of variance (classrooms by visits) revealed that the variation between classrooms in the incidence of children's interaction with an adult was significant (F=4·52, df 5, P<·01). The coefficient of reliability found by subtracting residual variance from classroom variance and expressing this answer as a proportion of classroom variance was 0·78 (Medley and Mitzel, 1963). Classroom variation also reached significance for children's interactions with other pupils (F=3·33, df 5, P<·05) and for the amount of time they were not interacting with anyone (F=3·96, df 5, P<·01). The associated coefficients of reliability for these two situations were 0·70 and 0·75 respectively.

TABLE 1

Classroom Variation in Children's Situations
(Average Classroom Percentages)

Situation	Classroom						
	A	B	C	D	E	F	Ave.
Interacting with adult (generally teacher)	7	17	13	8	8	16	12
Interacting with other pupil(s)	39	40	23	27	33	23	31
Not interacting	54	43	63	65	60	61	58

Interaction with Adults.

The frequencies of each item in category 1 were summed for the six lessons observed in each classroom and expressed as percentages of the total number of recorded pupil-adult contacts. The six classroom profiles so obtained were then averaged. The item frequencies in categories 2, 3a and 3b were processed in the same way. The average classroom profiles for all four pupil-adult categories are displayed in Table 2. The low incidence of pupil-observer interaction indicates that observers can be trained to adopt a relatively unobtrusive non-participant role in informal classrooms : in fact, children only watched or interacted with observers 12 times out of a total of 4,191 recorded observations.

TABLE 2

AVERAGE INCIDENCE OF ITEMS IN THE PUPIL-ADULT CATEGORIES
(Average Classroom Percentages)

Category	Item	Brief definition of item	Classroom Average
1. Target's Role	INIT	Target initiates to adult privately or publicly (Focus)	12
	STAR	Adult initiates or sustains contact with target privately or publicly (Focus)....	25
	PART	Target part of adult's group or class audience (Audience)	41
	LSWT	Target listens/watches adult interact privately or publicly with other child (ren) (Audience)	23
2. Interacting Adult	TCHR	Target interacts with teacher	97
	OBSR	Target interacts with observer	3
	OTHER	Target interacts with any other adult (head, secretary, etc.)	1
3a. Adult's Interaction	WORK	Adult interacts about a work matter	73
	ROUTINE	Adult interacts about a matter of classroom routine (organisation and management)	19
	POS	Adult reacts positively (praises)	1
	NEG	Adult reacts negatively (criticises)	7
3b. Adult's Communication Setting	IND ATT	Adult gives private individual attention to target	32
	GROUP	Adult interacts with group which includes target	18
	CLASS	Adult interacts with whole class	40
	OTHER	Adult gives private attention to another child or group and target listens in or eavesdrops	10

Interaction with Other Pupils.

The frequency of each item in category 4 was expressed as a percentage of the number of recorded pupil-pupil contacts for each lesson. The percentages for the six lessons observed in each classroom were averaged and then the average profile for all six classrooms was calculated. The item frequencies in categories 5, 6, 7a, 7b and 7c were processed in the same way. The average classroom profiles for all six pupil-pupil categories are shown in Table 3.

TABLE 3

AVERAGE INCIDENCE OF ITEMS IN THE PUPIL-PUPIL CATEGORIES
(Average Classroom Percentages)

Category	Item	Brief definition of item	Classroom Average
4. Target's Role	INIT	Target initiates contact with other pupil(s)	39
	COOP	Target co-operates by responding to an initiation or by sustaining contact	61
	IGN	Target ignores attempted initiation	0
5. Role of Other Pupil(s)	INIT	Other pupil initiates contact with target	25
	COOP	Other pupil co-operates by responding to target's initiation or by sustaining contact	75
	IGN	Other pupil ignores target's attempted initiation	0
6. Mode of Interaction	CNTC	Non-verbal, mediated solely by physical contact, gesture, etc.	4
	MTL	Non-verbal, mediated solely by materials, apparatus, etc.	10
	VRB	Verbal (may be accompanied by non-verbal interaction)	87
7a. Task of Other Pupil(s)	S TK	Same or similar to target's task	85
	D TK	Different to target's task	15
7b. Sex and Number of Other Pupil(s)	SS	Target interacts with one pupil of same sex	78
	OS	Target interacts with one pupil of opposite sex	16
	SV	Target interacts with several pupils of same sex	2
	BG	Target interacts with boy(s) and girl(s)	5
7c. Group of Other Pupil(s)	OWN GP	From target's own base group	86
	OTH GP	From another base group	14

When category 4 was cross-tabulated against category 5 it was found that on average only 37 per cent of the interaction was sustained in the sense that it was in progress at the previous time signal 25 seconds earlier. When category 7b was cross-tabulated against children's 'home' base a strong sex bias was detected. 76 per cent of the interaction was between children of the same sex, even when they belonged to mixed groups of boys and girls.

Activities and Locations.

The frequency of each item in category 8 was expressed as a percentage of the number of recorded events for each lesson. The percentages for the six lessons observed in each classroom were averaged and the average profile for all six classrooms was calculated. The item frequencies in categories 9a and 9b were processed in the same way. The average classroom profiles for the three categories which are applicable in all situations will be found in Table 4.

Table 4 groups the items in category 8 into four major types of activity: involved, non-involved, waiting for the teacher and a miscellaneous selection of other activities.

TABLE 4

AVERAGE INCIDENCE OF ITEMS IN THE CATEGORIES WHICH ARE APPLICABLE IN ALL SITUATIONS (Average Classroom Percentages)

Category	Item	Brief definition of item	Classroom Average
8. Target's Activity	COOP TK	Involved and co-operating on task work (e.g., reading)	55
	COOP R	Involved and co-operating on routine work (e.g., sharpening a pencil)	7
	MOB (WK)	Involved and mobile/work-oriented	5
	DSTR	Non-involved and distracted from work	10
	DSRP	Non-involved and disrupting another pupil's work	1
	HPLY	Non-involved and engaging in horse-play with other pupil(s)	1
	MOB (DSTR)	Non-involved and mobile/distracted	2
	WAIT TCHR	Waiting to see the teacher	4
	CODS	Partially distracted and partially co-operating on work	8
	DSTR WK	Distracted by another pupil's work	3
	DSTR TCHR	Distracted by teacher's activity or interaction with other pupil(s)	2
	RIS	Responding to internal stimuli (e.g., daydreaming)	2
	WOA	Working on a work activity which is not approved work	0
9a. Target's Location	P IN	Target in base group	80
	P OUT	Target out of base group	18
	P OUT RM	Target out of room	2
9b. Teacher Activity and Location	T PRES	Teacher physically present or interacting with target, his seated group or class	17
	T ELSE (I)	Teacher elsewhere interacting with other pupil(s)	80
	T APART (W)	Teacher apart and waiting/watching (not interacting)	2
	T REM (H)	Teacher removed and housekeeping (not interacting)	0
	T OUT RM	Teacher out of room (not interacting)	1

A two-way analysis of variance (classroom by visits) revealed significant variation between classrooms in the amount of involved activity ($F=5{\cdot}82$, df 5, $P<{\cdot}01$), non-involved activity ($F=3{\cdot}18$, df 5, $P<{\cdot}05$) and waiting for the teacher ($F=4{\cdot}73$, df 5, $P<{\cdot}01$). The associated coefficients of reliability were 0·83, 0·69 and 0·79 respectively.

When the four major types of children's activity were cross-tabulated against their two most common situations (interacting with other pupil(s) and not interacting), it was shown that children were less likely to be involved when interacting than when they were not. Full details are given in Table 5. The frequencies of each major type of activity were summed separately for the two situations in each classroom and then expressed as percentages of the total number of events falling into each situation during the six lessons observed. Spearman ' rho ' correlation coefficients showed positive relationships between the amount of involved activity in the two situations (0·26), between non-involved activity in the two situations (0·41) and between the overall amount of pupil-pupil interaction (Table 1) and the amount of involved activity when children were interacting (Table 5) (Spearman ' rho '=0·53).

TABLE 5

COMPARISON OF CHILDREN'S ACTIVITIES IN TWO SITUATIONS (INTERACTING WITH OTHER PUPILS AND NOT INTERACTING)
(Average Classroom Percentages for Each Situation)

Major Type of Activity	Situation	Classroom						
		A	B	C	D	E	F	Ave.
Involved	Interacting	51	57	44	52	32	46	47
	Not interacting	73	74	77	82	60	84	75
Non-involved	Interacting	27	23	31	23	41	24	28
	Not interacting	11	15	7	5	16	4	10
Waiting for the teacher	Interacting	1	0	1	0	3	0	1
	Not interacting	5	2	5	3	15	5	6
Other	Interacting	21	20	25	25	24	30	24
	Not interacting	10	8	11	10	9	7	9

DISCUSSION

The results of this exploratory study show that despite the relatively high amount of time children were interacting with other children (Table 1), away from their base place and mobile, they were involved with their work or waiting to see the teacher for almost three-quarters of each lesson (Table 4). In view of the relatively low amount of adult contact of any kind (Table 1) these findings support the common belief that children have a considerable ability to engage in independent or group work with a minimum of supervision.

However, the findings relating to pupil-pupil interaction suggest that it might be more difficult than is generally supposed to set up the group work conditions envisaged in the Plowden Report (1967) so that apathetic children " may be infected by the enthusiasm of a group " and able children benefit from " the cut and counterthrust of conversation . . . " (para. 757).

In the first place the observed sex bias militates against group discussions involving boys and girls, thereby reducing the chances for children to " learn to get along together, to help one another and realise their own strengths and weaknesses, as well as those of others " (para. 757). In the second place only half the interaction was concerned with the children's own work and it is debatable whether this incidence of involved activity is compatible with a view of group functioning in which children " get the chance of discussing, and so understanding more clearly, what their problem is " (para. 758). Finally, most of the recorded interaction had been initiated in the previous 25 seconds. This suggests that sustained conversations in which children explain and develop their ideas and arguments may be relatively uncommon and this casts some doubt on the extent to which children are gaining from " opportunities to teach as well as to learn " (para. 757).

Observation studies, such as the one reported here, enable checks to be made on commonly voiced assertions about the way children behave in different types of classroom. It is clear from this evidence that informal classrooms do not necessarily involve as much time wasting as their critics might imagine, nor as much sustained, work-orientated interaction as their advocates sometimes suppose.

ACKNOWLEDGMENT—This research forms part of Professor Brian Simon's SSRC funded project at the University of Leicester on ' The nature of classroom learning in primary schools.'

REFERENCES

BOYDELL, D. (1974). Teacher-pupil contact in junior classrooms. *Br. J. educ. Psychol.*, 44, 313-318.

MEDLEY, D. M. (1969). OScAR goes to nursery school : a new technique for recording pupil behavior. Paper presented at the meetings of the American Educational Research Association, Los Angeles, California.

MEDLEY, D. M., and MITZEL, H. E. (1963). Measuring classroom behaviour by systematic observation. In GAGE, N. L. (Ed.), *Handbook of Research on Teaching.* Chicago : Rand McNally.

MEDLEY, D. M., *et al.* (1973). The Personal Record of School Experiences (PROSE). In BOYER, E. G., SIMON, A., and KARAFIN, G. (Eds.). *Measures of Maturation : An Anthology of Early Childhood Observation Instruments, Vol. II.* Philadelphia : Research for Better Schools, Inc.

NIE, N. H., BENT, D. H., and HULL, C. H. (1970). *The Statistical Package for the Social Sciences.* New York : McGraw-Hill.

PLOWDEN REPORT (1967). *Children and their Primary Schools.* London : HMSO.

SIMON, A., and BOYER, E. G. (1970). *Mirrors for Behavior II : An Anthology of Observation Instruments.* Philadelphia : Research for Better Schools, Inc.

(*Manuscript received* 20*th August*, 1974)

Br. J. educ. Psychol., 45, 130-140, 1975

DECISION RULES FOR TEACHING STRATEGIES IN PRIMARY SCHOOLS : PERSONALITY-TREATMENT INTERACTIONS

By E. ANNE TROWN
(*North Riding College of Education*)*

AND G. O. M. LEITH
(*Department for Educational Research and Development, University of Utrecht*)

SUMMARY. A mathematical task was used among 432 junior school pupils in a search for personality-teaching strategy interactions. Factors of sex, ability, anxiety, extraversion, model and teaching strategy were incorporated into a replicated design. An inductive, learner-centred exploratory strategy and a deductive teacher-centred supportive strategy each employed the same mathematical models. There was no evidence of strategy-ability or strategy-extraversion interactions but the strategy-anxiety interaction was significant. Linear functions, relating the retained learning outcome to anxiety score, were used as the basis of decision rules for differential instruction. Pupils at the lower end of the anxiety range may be assigned to the exploratory strategy and those at the upper end of the range to the supportive strategy if advantage is taken of the disordinal nature of the observed interaction. The consistency of a negative linear relationship between outcome and anxiety for the exploratory strategy contrasted with the complex pattern associated with the supportive strategy. In the latter case, linear functions of differing sign afforded predictability for subgroups of introverts and extraverts and for subgroups of low ability and high ability children. When all data were combined for the supportive strategy, however, a non-linear relationship was revealed.

INTRODUCTION

IF a line of regression relating learner-aptitude to learning outcome is drawn for each treatment under consideration, then non-parallelism will indicate the presence of an aptitude-treatment interaction (ATI). An ordinal interaction is defined as one in which the order of treatments A and B is the same over the entire aptitude range, i.e., one treatment is superior overall (Figure 1). Where the regression line for treatment A crosses the regression line for treatment B within the aptitude range, however, the interaction is said to be disordinal and decision rules for differential instruction may be formulated. Pupils at one end of the aptitude range may be assigned with advantage to one treatment, while those at the other end of the range are assigned to the alternative treatment.

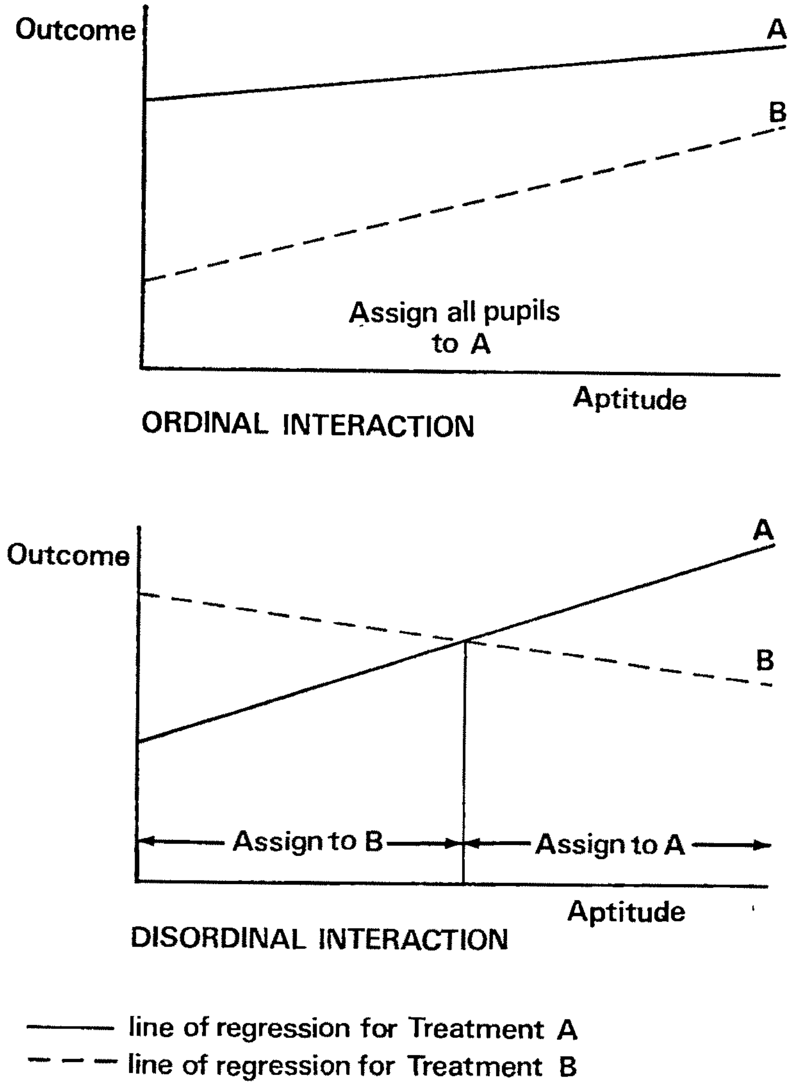
Where analysis of variance is employed then mean outcome scores will be examined at blocked high and low levels of aptitude. A significant F-ratio for interaction, with crossing graphs of cell means, was formerly considered sufficient for disordinality (Lindquist, 1953 ; Lubin, 1961), but Bracht (Bracht and Glass, 1968) proposed more stringent criteria, conservative in relation to the risk of false rejection of the ordinality hypothesis. Differences between outcome means at both high and low levels of the aptitude variable were required to be separately significant, as well as opposite in sign. Reaction against the severity of this suggestion (Leith, 1972 ; Cronbach and Snow, 1973) has been based upon the extremely low probability of joint occurrence of two separately significant events. A recent distinction has been made between research set up in the course of the development of a theoretical model and studies which are concerned with rules for differential instruction. Plomp (1973) and Gustafsson (1974) have suggested that a significant F-ratio for interaction, together with the simple crossing of regression lines, should be considered sufficient for disordinality in the former, but not in the latter, case.

---

* Dr. Trown is now in the Department of Educational Research, University of Lancaster.

FIGURE 1

APTITUDE-TREATMENT INTERACTIONS
PATTERNS OF LINES OF REGRESSION



Gage and Unruh (1967), Bracht (1970) and Cronbach and Snow (1969) have drawn attention to the scarcity of well established, educationally significant aptitude-treatment interactions. An inconsistency of pattern arises from specific studies taken over the whole field and much of the evidence arising from the initial period of research is equivocal in nature. In a systematic search among 90 studies designed to reveal possible aptitude-treatment interactions, Bracht (1970) found only five to meet his own proposed criteria. These five experiments were characterised by controlled presentation of treatment and by the non-complex nature of the learning material. Three rested upon personality variables, included by Bracht among types of aptitude variable likely to give rise to disordinal interactions by virtue of their factorial simplicity. A survey of personality research by Cronbach and Snow (1969) revealed a disappointing position in terms of useful and consistent interactions. Difficulties were attributed to the interdependence of multiple aspects of personality and to the lack of adequate definition of teaching variables.

On a mathematical task, it would perhaps seem likely that the best matching of teaching strategies to children would be through some kind of ability measure. Measures of general ability are considered by Bracht to be unlikely to give rise to disordinal interactions because of their high correlation with performance on most complex tasks. Cronbach and Snow (1969), on the other hand, conclude that general ability, though often vaguely defined, does enter into interactions but they draw attention to the few results showing the existence of expected effects in relation to spatial and other specialised abilities. In a previous experiment among 12-year-olds learning about vectors, the present writer found some non-significant interaction of teaching method with level of general ability, but there was a more consistent and significant interaction with level of extraversion (Trown, 1969, 1970 ; Leith and Trown, 1970). Among younger children, interactions between teaching method and anxiety have been reported by Leith and Bossett (1967) and by Leith (1973). Closely structured sequences, incorporating formal guidance, have favoured anxious children while absence of structure has aided the non-anxious.

LEARNING MATERIAL

The chosen topic is described in the introduction to the Nuffield Teachers' Guide *Computation and Structure* 4 (Nuffield Mathematics Project, 1969) . . . " an attempt has been made to introduce (integers) on a sound footing (so that the children can really understand, without being given ' tricks ') but within the ' primary ' framework." Learning material, based upon the ideas set out in this guide, was written by the experimenter and developed during a one-year trial period in Leicestershire schools. It was designed to introduce the set of directed integers, $\{. . . {}^-3, {}^-2, {}^-1, 0, {}^+1, {}^+2, {}^+3, . . .\}$ and to investigate their behaviour under the operation of addition.

A directed integer may be defined as a set of ordered pairs, e.g., ${}^-2$ may be defined as $\{(0, 2), (1, 3), (2, 4), (3, 5), . . . \}$. Each of these ordered pairs may then be represented physically in many ways. For instance, the ordered pair (1, 3) may be represented by counters on a board (one counter on the left-hand side, three on the right-hand side), by means of washers hung on a mathematical balance or equaliser (at ' position 1 ' on the left-hand side and ' position 3 ' on the right-hand side), by means of the lengths of a pair of rods, etc. With the mathematical balance it is particularly easy for the child to recognise the equivalence of number pairs. If a pair of washers are hung at the back of the equaliser by the teacher, then the child may make a long list of the ways in which he is able to restore balance, using only two washers at the front of the equaliser. In doing so, the pupil sets up the equivalence class of ordered number pairs defining some integer. The addition of two directed integers involves the combination of two such separate effects. This learning material, while especially designed for the age-range, was selected as new and potentially 'stretching' for the pupils concerned. There has been some suggestion that increase in item difficulty (Leith *et al.*, 1970) or increase in pacing schedule (Firmin and Webb, 1971) may accentuate differences in the way in which personality affects learning.

In the case of the supportive strategy, the sequence employed, over each of the 12 sections of learning material, was that of teacher-provided statement of organising principle, followed by related pupil activity with mathematical models, and subsequent restatement of principle by the teacher. Such statements were both spoken, and written on the blackboard. The same activities with models were used in the exploratory strategy, but this time at the beginning of each section of the learning sequence. Each statement of principle by the teacher was now delayed until pupils had been given the opportunity to perceive the relationship themselves and had been encouraged to attempt an appropriate generalisation.

The ideas of Bruner (1964) were drawn upon in establishing a system of five mathematical models, to be used in association with each of the teaching strategies. Four of the models were expressed in concrete or diagrammatic mode but the fifth employed written pairs of number symbols throughout. A fuller description of learning material, models and strategies may be found elsewhere (Trown, 1973).

The teaching team of the Mathematics Research Project of the University of Hull (Land *et al.*, 1966, 1967, 1968, 1969) have noted the reluctance of class teachers to persist with a strategy that is apparently proving ineffective among their own pupils. It was decided, on this occasion, to accept the limitations that are involved in the use of one experimenter, even one with varied and substantial experience in the teaching of mathematics, as the only 'teacher' used to control treatment. A complete pattern of instruction, including principles, definitions, questions and use of blackboard, was specified for each teaching strategy. Each child was provided with pencil and paper and individual pieces of mathematical equipment, where appropriate. Management of this apparatus was pre-arranged.

Since there was no desire to suppress the 'noise' of the classroom, children's incidental comments, and any responses to them, were always spontaneous. Because of the planned sequences and interplay of strategy and model, however, these happenings were always in the nature of variations upon a controlled pattern.

PROCEDURE

The 215 boys and 217 girls who took part formed the third and fourth year classes of four junior schools in North Yorkshire. The mean age of the boys, and of the girls, was 10 years 6 months. The idea of assembling, in some central location, randomly chosen groups from many junior schools, was rejected as too far removed from the usual classroom situation. Each school, chosen for variety of pupil background and of teaching method, had previously employed a mixture of formal and informal approaches to the teaching of mathematics. No attempt was made to reduce school centred sources of variance in the subsequent statistical analysis.

Where mathematical models are being used, primary school classes are often organised into small groups. It was decided, therefore, that within the experiment no more than 10 children should be taught on any one occasion. In each school a small room was used for the experiment so that each lesson could take place without outside interruption. The mean time for treatments using the supportive strategy was 32 minutes with a corresponding mean time for the exploratory strategy of 35 minutes.

A non-verbal test (Raven's Progressive Matrices, Raven, 1960) was used to measure ability. The HB Personality Inventory (Hallworth, 1962), which had been designed as a short, easily administered, test foı use with children, provided orthogonal measures of anxiety and extraversion, 18 items relating to anxiety and 10 to extraversion.

Dichotomies, about median scores for ability, anxiety and extraversion, were undertaken, so that, of the 32 children randomly assigned to each of the 10 treatments, two of each sex were of above average ability, of more than average anxiety, and of more than average extraversion, and so on for all possible combinations of levels of these characteristics. The 112 children who remained were assigned to the 'control' group. These pupils completed all tests but were given no treatment. Treatments were administered in random order in each school.

The retention test (used also as a pre-test) was on a prepared sheet with blanks left for pupil response. It tested the ability to perform the following operations.

(*a*) to select an ordered pair from the equivalence class defining a directed integer ;

(*b*) to add two ordered pairs ;

(*c*) to identify the directed integer represented by the resultant ordered pair ;

(*d*) by combining these processes, to add two directed integers ;
e.g., $^{+}2+^{-}3=(3,1)+(1,4)=(4,5)=^{-}1$.

The pre-test was given in each school at the beginning of the experimental period, and the retention test either 17 or 18 days after the end of that period, according to school. Test sheets required symbolic response and no concrete or diagrammatic aid was available to children when they were tested. Maximum score was 40.

One source of controversy in ATI studies lies in the possible use of gain scores, or of raw post-test scores, as a measure of learning. In this experiment pre-test scores were low (Table 2) and so differences between gain scores and raw retention-test scores were not great. Unadjusted gain scores (pre-test to retention test gain) were used as a measure of retained learning.

RESULTS

The teaching strategies were evidently well matched overall, the difference in mean retained learning score between the deductively organised supportive strategy (16·40) and the inductively organised exploratory stıategy (16·30) being minimal. Two pieces of missing retention test data for the experimental group were replaced, conservatively, by the mean score of all other subjects on the same measure.

TABLE 1

MEAN SCORES ON PRE-TEST AND RETENTION TEST (BY SEX).

		Experimental Group			Control Group		
		Boys	Girls	All	Boys	Girls	All
Pre-Test ..	(N) .... Mean........	(160) 2·10	(160) 1·79	(320) 1·94	(55) 2·02	(57) 1·19	(112) 1·60
Retention-Test	(N) .... Mean...... ..	(158) 18·06	(160) 18·51	(318) 18·29	(53) 4·98	(56) 3·96	(109) 4·46

A six-factor analysis of variance was carried out on the retained learning (pre-retention gain) data. For models, strategies and sex as main effects, F-ratios were each less than 1.

The following main effects were significant :

(*a*) An ability effect significant at the 0·1 per cent level.

(*b*) An anxiety effect (non-anxious scoring more than anxious) significant at the 2½ per cent level.

Neither the strategy-extraversion interaction nor the strategy-ability interaction produced an F-ratio greater than 1. The strategy-anxiety interaction, however, was significant at the 2½ per cent level (F=6·51, 1 and 158 df.) In line with previous research among younger pupils, it was evident that any strategy-personality interaction arising from the present data concerned levels of anxiety rather than of extraversion.

TABLE 2

STRATEGY-ANXIETY INTERACTION (BY SEX). MEAN SCORES FOR RETAINED LEARNING.

	Boys (N per cell=40)		Girls (N per cell=40)		All Pupils (N per cell=80)	
	Low Anx.	High Anx.	Low Anx.	High Anx.	Low Anx.	High Anx.
Supportive Strategy	15·98	15·26	16·63	17·73	16·30	16·49
Exploratory Strategy	20·30	12·33	18·30	14·25	19·30	13·29

Anxiety level distinguished between those who were able to profit greatly from the learner-centred exploratory approach and those whom it clearly handicapped (Table 2). The teacher-centred supportive strategy, on the other hand, was almost equally effective at each level of anxiety.

Even though differences were in the direction indicated by previous research, a conservative approach was adopted in the use of two-tailed t-tests at each anxiety level. Consequently, the interaction for all pupils cannot be said to meet the most severe of disordinality criteria (significant differences between strategy-outcomes at each level of anxiety). Nevertheless the t-tests showed that differences as large as each of those observed would arise, by chance, on less than one occasion in 10. The probability of simultaneous occurrence of two such differences, in opposite directions, by chance, is less than ·005.

TABLE 3

STRATEGY-ANXIETY INTERACTION (BY LEVELS OF EXTRAVERSION AND ABILITY). MEAN SCORES FOR RETAINED LEARNING.

(i) Extraverts (N per cell=40)			(ii) Introverts (N per cell=40)		
	Low Anx.	High Anx.		Low Anx.	High Anx.
Supportive Strategy ..	17·48	13·58	Supportive Strategy ..	15·13	19·40
Exploratory Strategy ..	19·18	10·80	Exploratory Strategy ..	19·43	15·78
(iii) High Ability (N per cell=40)			(iv) Low Ability (N per cell=40)		
Supportive Strategy ..	22·75	19·63	Supportive Strategy ..	9·85	13·36
Exploratory Strategy ..	26·23	17·85	Exploratory Strategy ..	12·38	8·73

The strength of the strategy-anxiety interaction was demonstrated by its persistence over subgroups of introverts and extraverts and subgroups of low and high ability pupils (Table 3). In the subsequent regression analysis, each child's anxiety score was considered in relation to his eventual retained learning score. Regression lines over the entire anxiety range were drawn for each subgroup of children. In the case of the exploratory strategy, every regression line was negative in gradient (Table 4). For every subgroup, anxiety was debilitating. When all data were combined for the exploratory strategy, a hypothesis of zero linear regression was rejected at the 2½ per cent level.

TABLE 4

STRATEGY-ANXIETY INTERACTION : GRADIENTS OF LINES OF REGRESSION.

	Introverts	Extraverts	Low Ability	High Ability
Supportive Strategy	+0·42	−0·44	+0·31	−0·31
Exploratory Strategy..	−0·38	−1·00	−0·44	−0·74

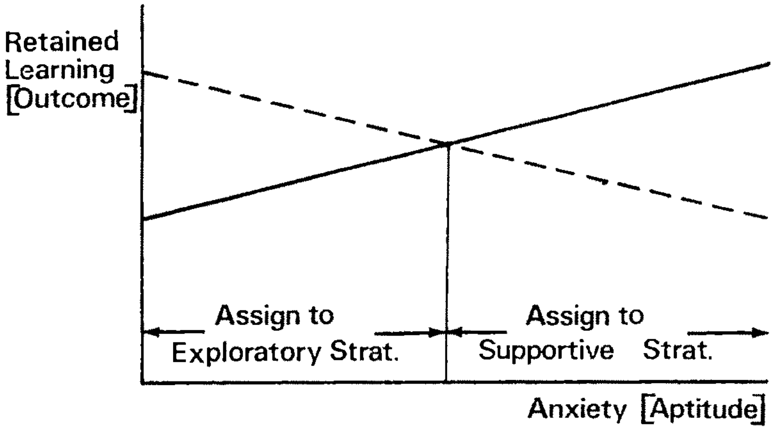
In the case of the supportive strategy, regression lines afforded predictability of retained learning outcome on a basis of anxiety scorefor subgroups of children. These lines were positive in gradient for introverts and less able children (i.e., increase in anxiety score was actually helpful), but negative in gradient for extraverts and more able children (Table 4). Two different patterns of crossing regression lines emerged therefore, the first among the introverts and less able, and the second among the extraverts and more able (Figure 2). Both patterns were disordinal, regression lines crossing within the anxiety range in each case. Decision rules were formulated for each subgroup. These rules indicated the anxiety score at which choice of strategy should be reversed (determined by the position of the point of intersection of the regression lines). (Table 5.)

The amount of advantage to be forecast by assignment to 'right' rather than 'wrong' strategy may be illustrated by the predicted mean retained learning outcome (18·98 out of 40) for non-anxious introverts assigned by decision rule to the exploratory strategy compared with the predicted mean outcome (15·75 out of 40) if the same group of pupils were to be assigned to the supportive strategy.

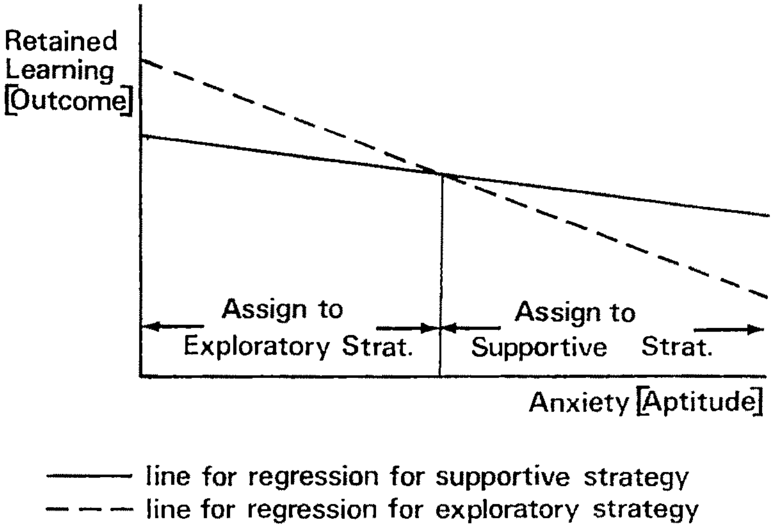
FIGURE 2

STRATEGY-ANXIETY INTERACTION
PATTERNS OF LINES OF REGRESSION (not to scale)

FIRST DISORDINAL PATTERN



SECOND DISORDINAL PATTERN



In the case of the supportive strategy, the indications were that if all children were considered together then subsets of data, exhibiting linear relationships of differing sign, would combine to show an overall non-linear relationship. This was confirmed when gradients, and F-ratios for gradient, were calculated on the basis of linear regression for all supportive strategy data. F-ratios were very small, and so was the proportional reduction in variance given a linear rule. It was decided, for this strategy, to examine means of low, medium and high anxiety groups, in order to check for direction of curvilinearity. All pupils with an anxiety score of 0 to 6 were placed in a low anxiety group, those with a score of 7 to 10 in an intermediate group, and those with a score of 11 to 18 in a high anxiety group. Numbers in cells were not large, but there was a clear indication that, for the supportive strategy, the overall relationship between outcome and anxiety was basically curvilinear (' inverted-U ') in form, with an intermediate level of anxiety optimal for both boys and girls (Table 6).

TABLE 5

Strategy-Anxiety Interaction : Decision Rules for Teaching Strategies.

Decision	Introverts	Extraverts	Low Ability	High Ability
Assign to Exploratory Strat.	Anxiety scores between 0 and 9	Anxiety scores between 0 and 6	Anxiety scores between 0 and 7	Anxiety scores between 0 and 9
Assign to Supportive Strat. ..	Anxiety scores between 10 and 18	Anxiety scores between 7 and 18	Anxiety scores between 8 and 18	Anxiety scores between 10 and 18

(Maximum anxiety score was 18.)

TABLE 6

Mean Retained Learning Scores for Low/Medium/High Anxiety Groups (By Sex) for the Supportive strategy.

	Low Anxiety		Medium Anxiety		High Anxiety	
	(N)	Mean	(N)	Mean	(N)	Mean
Boys	(31)	16·4	(24)	17·3	(25)	13·1
Girls	(22)	16·2	(29)	17·8	(29)	17·3
All	(53)	16·3	(53)	17·5	(54)	15·4

DISCUSSION

Curriculum reformers in primary mathematics have placed their emphasis as much upon how to learn as upon what to teach. The Nuffield Mathematics team themselves make the point in the general introduction to their guides. " Running through all the work is the central notion that the children must be set free to make their own discoveries and think for themselves, and so achieve understanding, instead of learning off mysterious drills." The questions of motivation, and of changes in pupil attitude toward mathematics, are raised and deserve investigation in their own right. It has never, however, been conclusively proved that ' understanding ' is best achieved through the use of any particular teaching method.

The results of the present experiment suggest that some, but by no means all, children remember more when they are ' set free ' to learn. A teacher could plan to take advantage of the observed interaction by setting up exploratory situations in the classroom, but at the same time taking care to provide anxious pupils with the kind of support associated with more traditional teaching methods. The interaction is a strong one, consistent over sub-groups and backing up previous research. The question of transferability to other learning situations still remains, however, for the history of aptitude-treatment interaction research is littered with confusing results and failures to replicate. There is a clear need to set up studies which are seen to be relevant to the teacher's own decision making. Evidence relating to younger children would be particularly helpful to primary school teachers, who already organise group learning situations with the needs of particular children in mind.

Treatments in British studies have rarely been designed to compensate for learner deficiencies or to exploit learner abilities and most ATI research has been carried out in the USA. Webster and Mendro (1974) point to the slowness of educational evaluators in incorporating the results into decision processes. They conclude that " obsession with main effects may have resulted in the loss of much crucial information needed to support critical educational decisions " (p. 86). Nor can the preoccupation of researchers with clarification of concepts and establishment of criteria have helped the processes of classroom implementation.

Biggs (1962), after his long-term investigation into the learning of primary school arithmetic, states of the relationship between anxiety and achievement : "Although curvilinear functions were occasionally obtained, the most frequently encountered and the most meaningful functions were linear . . . Nevertheless, the treatment demonstrated the main point, namely, that the action of anxiety is not uniformly positive or negative and that whether it is positive or negative will depend upon further factors in the total learning-performance situation " (p. 348).

This main point is again demonstrated in the present experiment. The overall pattern over strategies provides evidence, scarce in classroom learning situations, which may be interpreted as supporting the Yerkes-Dodson Law (Yerkes and Dodson, 1908). This law may be used to predict that in ' easier ' complex learning situations (perhaps to be equated with the supportive strategy) an intermediate level of anxiety will be optimal. In ' more difficult ' situations (perhaps akin to the exploratory strategy), however, the optimal level of anxiety will be lower and the relationship basically a negative linear one.

Analysis over other dimensions of pupil personality and ability is an aid to fuller insight into the strategy-anxiety interaction. Especially did this prove to be the case in the present study, when linear functions were sought as a basis for teaching-strategy decision-rules. Complex inter-relationships may result in apparently conflicting results, but they also offer the possibility of maximizing learning outcomes.

ACKNOWLEDGMENTS—This experiment formed part of the work undertaken for the D.Phil. degree of the University of Sussex. Thanks are due to Professor H. J. Butcher (co-supervisor) and to the head teachers, staff and pupils of Braeburn, Gladstone Road, Overdale and Wheatcroft Junior Schools, Scarborough, and of Woodstock Junior School, Leicester.

REFERENCES

BIGGS, J. B. (1962). *Anxiety, Motivation and Primary School Mathematics.* Occasional Publication No. 7. Slough : NFER.

BRACHT, G. H. (1970). Experimental factors related to aptitude-treatment interactions. *Rev. educ. Res.*, 40, 5, 627-645.

BRACHT, G. H., and GLASS, G. V. (1968). The external validity of experiments. *Am. educ. Res. Journ.*, 5, 437-474.

BRUNER, J. S. (1964). The course of cognitive growth. *Am. Psychol.*, 19, 1-15.

CRONBACH, L. J., and SNOW, R. E. (1969). *Individual Differences in Learning Ability as a Function of Instructional Variables.* Research report, Stanford University, California.

CRONBACH, L. J., and SNOW, R. E. (1973). *Aptitudes and Instructional Methods.* Stanford University, California.

FIRMIN, S. M., and WEBB, C. C. (1971). Pacing schedules, personality type and learning performance. In CLEARY, A., PACKHAM, D., and MAYES, T. (Eds.), *Aspects of Educational Technology, V.* London : Pitman.

GAGE, N. L., and UNRUH, W. R. (1967). Theoretical formulations for research on teaching. *Rev. educ. Res.*, 37, 358-370.

GUSTAFSSON, J. E. (1974). Verbal *versus* figural in aptitude-treatment interactions. Report No. 36, Institute of Education, University of Göteborg.

HALLWORTH, H. J. (1962). H.B. Personality Inventory. Unpublished ; available through University of Birmingham.

LAND, F. W., BISHOP, A. J., LEVY, L. B., and ROBERTSON, A. H. (1966, 1967, 1968 1969). *Annual Reports of the Mathematics Teaching Research Project.* University of Hull Institute of Education.

LEITH, G. O. M. (1972). *Personality, Intellectual Style and Study Background.* Research report. University of Utrecht.

LEITH, G. O. M. (1973). *The Influence of Anxiety on Achievement.* Research report. University of Utrecht.

LEITH, G. O. M., and BOSSETT, R. (1967). Mode of learning and personality. Research report, No. 14. National Centre of Programmed Learning. University of Birmingham.

LEITH, G. O. M., LISTER, A., TEALL, C., and BELLINGHAM, J. (1970). An evaluation of training systems for teaching conversion to decimal currency and exchange in mixed coinage. *Prog. Learn. educ. Tech.*, 7, 140-150.

LEITH, G. O. M., and TROWN, E. A. (1970). The influence of personality and task conditions on learning and transfer. *Prog. Learn. educ. Tech.*, 7, 181-188.

LINDQUIST, E. F. (1953). *Design and Analysis of Experiments in Psychology and Education.* New York : Houghton Mifflin.

LUBIN, A. (1961). The interpretation of significant interaction. *Educ. psychol. Meas.*, 21, 807-817.

NUFFIELD MATHEMATICS PROJECT (1969). *Computation and Structure* 4. The Nuffield Foundation, Edinburgh and London : Chambers and Murray. (Draft version, 1967.)

PLOMP, T. (1973). The statistical basis for ATI research : definitions and techniques. Paper presented at the Congress on Methodology of Research in Higher Education, Rotterdam.

RAVEN, J. C. (1960). *Standard Progressive Matrices.* London : H. K. Lewis.

TROWN, E. A. (1969). An experimental study of the effects of ten different teaching strategies in relation to the individual characteristics of anxiety, extraversion-introversion and intelligence. Unpublished M.Ed. thesis, University of Birmingham.

TROWN, E. A. (1970). Some evidence on the interaction between teaching strategy and personality. *Br. J. educ. Psychol.*, 40, 209-211.

TROWN, E. A. (1973). An investigation of structured approaches to the teaching of mathematics in relation to dimensions of pupil personality. Unpublished D.Phil. thesis, University of Sussex.

WEBSTER, W. J., and MENDRO, R. L. (1974). The investigation of aptitude-treatment interactions as an integral part of programme evaluation. *J. Exp. Educ.*, 43, 86-91.

YERKES, R. M., and DODSON, J. D. (1908). The relation of strength and stimulus to rapidity of habit formation. *J. comp. Neurol. Psychol.*, 18, 459-482.

(*Manuscript received 5th September,* 1974)

Br. J. educ. Psychol., 45, 141-152, 1975

BOREDOM AT SCHOOL

By W. P. ROBINSON

(*School of Education, Macquarie University, New South Wales*)

SUMMARY. The absence of theories of boredom is argued to be a gap both in psychology and education. A model of boredom is put forward as a basis for specifying predisposing factors and possible consequences. For the child in school, the model has the character of a positive feedback loop that is maintained in unstable equilibrium by external forces until such time as he can leave school. A secondary analysis of data from the national sample of Young School Leavers (Morton-Williams and Finch, 1968) provides support for almost all the hypotheses tested.

INTRODUCTION

BOREDOM is a neglected topic in both psychology and education. In his treatment of conflict, arousal and curiosity, Berlyne (1960) refers to boredom as a drive state whose onset is provoked when 'external stimuli' are excessively scarce or excessively monotonous (p. 187) and whose occurrence leads to 'divertive exploration.' There seems to have been no development of this.

Government reports on education make frequent mention of curiosity and boredom and presume positive associations between educationally desirable objectives and curiosity. While reports on primary education (Plowden, 1967) are sanguine about the extent to which intrinsic motivation is being fostered, those on secondary education make more frequent reference to problems of boredom. "Too many appear to be bored and apathetic in school . . ." (Newsom, 1963, 47). Both Crowther, (1959) and Newsom noted a high incidence of 'early leaving' and suggested that a perceived lack of usefulness and relevance in the curriculum was, in large measure, responsible for this. The only empirical study exploring the nature and extent of this association has been the Schools Council inquiry into 'early leaving' (Morton-Williams and Finch, 1968).

In spite of its intrinsic interest for psychologists and its practical significance for educators, boredom remains otherwise unexplored. Through the efficient and effective helpfulness of Moreton-Williams and Finch, it became possible to conduct a secondary analysis of their data and thereby begin to open up this field of inquiry. There is no pretence to answer too many questions, but conceptual worries can be posed, a model for operation of boredom in secondary schools proposed, and its potential utility subjected to initial probing.

For Berlyne, 'boredom' results from a pattern of stimulation that allows a physiological state of high internal arousal to develop. He argues for 'high' arousal because restlessness, agitation and emotional upset are commonly mentioned symptoms, but this analysis may be too simple and restrictive. In everyday discourse "He was getting bored, so he leapt up to take the dog for a walk" and "He fell asleep because he was bored" are as permissible as "He was getting bored so he started throwing paper darts." The difference between the three appear to reside not in the state but in the responses preferred, possible and permitted. If a person has coping strategies and is free to adopt them, he can escape from incipient boredom. One sedentary set of such strategies is daydreaming or falling asleep. But if external constraints do not permit either of these, the becoming bored person can struggle to maintain interest or engage in diversionary activities. He is liable to become restless, agitated and emotionally upset.

If these distinctions are sensible, then the application of the word ' boredom' to each is understandable, but could confuse. An emphasis upon the similarity of the inducive stimuli encourages the adoption of a single term, a focus upon the differences in the reactive behaviour leads to a worry that this is inappropriate. It is presumably the combination of incipient boredom in conflict with the teacher's demands that is the common classroom problem. An immediate image is of a teacher droning on—with an eagle eye alert for wandering attention. The threat of immediate sanctions, school norms, and finally the law of the land, dictate that the pupils stay where they are, that they attend and learn. Teachers are under no correlative obligation to sell the syllabus, only to cover it.

The existence and use of sanctions introduces complications into a conceptual analysis of boredom. It appears to be the case that sometimes children say they are bored when they mean they are anxious ; they will say a subject bores them when they, in fact, are frightened of the consequences of failure. We have to ignore this complication and with our data we have to treat reported boredom as a unitary phenomenon, while allowing it to have a diversity of associated responses. At some later point in time, it may well prove necessary to complicate and improve the conceptual analysis.

But are pupils bored ? If so, who is most bored by what ? What are the antecedents and consequents of being bored ? Moreton-Williams and Finch (1968) were able to estimate the incidence of interest in individual school subjects (pp. 59 and 74), to point out the nature of the covariation between judged uselessness and boringness for various subjects (p. 60), and to comment on what made subjects boring (p. 79) ; the same thing all the time, teachers going on and on, and lack of variety were mentioned by over 40 per cent of pupils, and a similar number referred to ' not understanding,' ' not being any good at,' and ' not having subjects explained enough.'

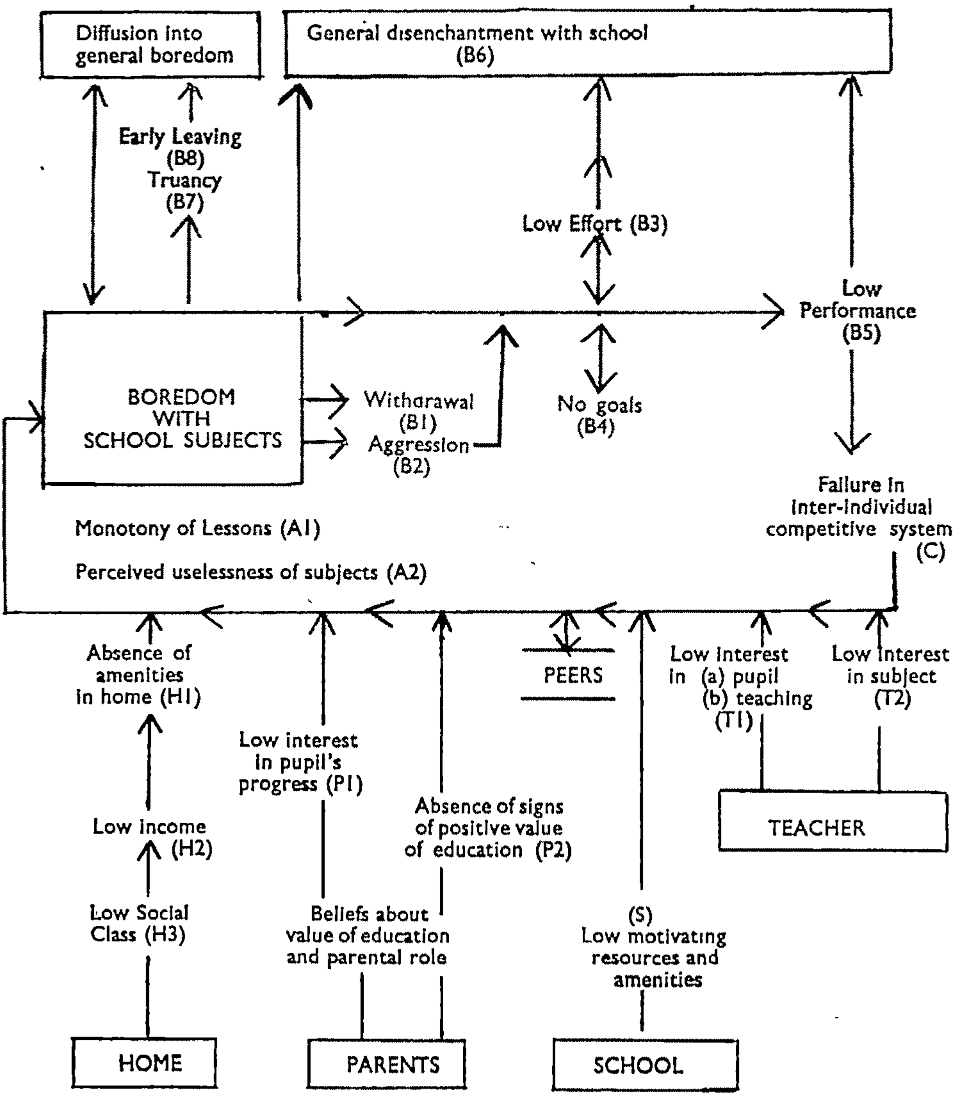
In our analysis we sought to specify some possible antecedents and consequents of boredom, initially developing a model and hypotheses constrained by the nature of the data we thought could be available for analysis (see Figure 1).

The model should not be assumed to be complete in content or form. The influences of peers and their pressures and of schools are not developed because we had no relevant data. The actual directions of influence are matters for experimentation rather than correlation. The model should *not* be construed as a tight causal sequence, such that, for example, the presence of factor A is treated as either a necessary or sufficient condition of the inducement of boredom. The arrows are labelled ' increased the likelihood,' and perhaps this should be qualified further by adding ' given present conditions in schools.' For example, the probability of a pupil making little effort in his study of a subject will be affected by factors other than 'boredom,' but, other things being equal, if he is bored he is not so likely to make an effort than if he is interested. Boredom will also have more means of expression than are listed. Which means of coping with it are generally adopted by pupils will probably be a function of least effort and minimum cost principles, interacting with past experience and present perception. The hypotheses offered are not predictions derived from a theory, but are first steps towards an organisation of the facts that need explanation and should be evaluated as such. Hypotheses actually examined are marked with a letter and number.

The central feature is ' boredom with school subjects.' The first question that might be asked is whether scores on other questions relating to boredom with school subjects are associated with the general index. Possible immediate antecedents are experienced regularity (A1) and perceived uselessness of the subjects studied (A2). Immediate expressions of boredom might be withdrawal into self (B1), aggression (B2) (diversionary activities) and/or reduced effort in learning boring subjects (B3).

FIGURE 1

REPRESENTATION OF DYNAMIC OF BOREDOM AND OTHER VARIABLES.


Diffusion into general boredom
General disenchantment with school (B6)
Early Leaving (B8)
Truancy (B7)
Low Effort (B3)
Low Performance (B5)
BOREDOM WITH SCHOOL SUBJECTS
Withdrawal (B1)
Aggression (B2)
No goals (B4)
Failure in inter-individual competitive system (C)
Monotony of Lessons (A1)
Perceived uselessness of subjects (A2)
Absence of amenities in home (H1)
Low interest in pupil's progress (P1)
PEERS
Low interest in (a) pupil (b) teaching (T1)
Low interest in subject (T2)
Low income (H2)
Absence of signs of positive value of education (P2)
TEACHER
Low Social Class (H3)
Beliefs about value of education and parental role
(S) Low motivating resources and amenities
HOME
PARENTS
SCHOOL

Reduced effort might well be associated with a lowering of or an abandonment of performance-centred aspirations for improvements (B4), and either or both should lead to lower performance (B5) (Atkinson, 1964). In the absence of extra-learning attractions at school, truancy is one intermittent way of avoiding boredom (B7), while leaving school as soon as possible finalises escape (B8). One might expect boredom with school subjects to generalise to a general disenchantment with school (B6) and to increase the chances of developing a chronic state of general boredom (B9), while recognising the probability of influence in the reverse direction as well.

The consequences of low performance are likely to be contingent upon the values placed upon academic success by school, parents and peers, but, other things being equal, one might expect feelings of failure to be more pronounced when the school is formally competitive. Streaming and a high incidence of pupils taking national examinations are two symptoms of competition (C). A chronic history of failure and an expectation of future failure, it has been suggested, may in turn increase boredom or anxiety states reported as boredom. It should be noted that the system is a closed cycle and, once activated, will tend to sustain itself. However, external influences can also affect it.

Pupils can use teachers as models and models influence behaviour (Bandura and Walters, 1963) . If teachers are bored by their subjects (T2) (no data), or by the teaching of their subjects (no data), pupils should be less likely to be enthusiastic about them. Teachers can also show no respect for or interest in either the learning of their pupils or the pupils as people (T1). Lack of either example or encouragement from teachers should tend to maintain or increase boredom. The behaviour of teachers will, of course, be constrained by the resources of the school (almost no data) and the behaviour of other teachers and heads—and so on up to the Minister of Education and Science. Both the interest and encouragement offered by peers (no data) and parents (P) should affect boredom, and these in turn will be expressions of more general beliefs about the value of education for living. As with teachers, the power of parents and peers as models should not be forgotten.

Children can lack material resources that might prevent or ameliorate the boredom—low performance cycle, for example, overcrowding in the home (H1). These amenities will be in part, but only in part, contingent upon the wealth and income of families (H2). The concept of ' social class ' normally embraces many covarying features of different life-styles that should affect the boredom of children with school subjects. Children from lower working class backgrounds should be particular victims of the boredom cycle (H3).

METHOD

Full details of the sampling design and specific measures used are to be found in Moreton-Williams and Finch (1968) or Robinson (1974).

Subjects.

The sample contained 4,618 13 to 16-year-old children. The record of one boy had been nibbled away by a mouse, leaving 4,617 subjects. " The sample design was a two stage multi-stratified random one. At the first stage, a sample of maintained secondary schools in England and Wales was taken; at the second samples of 13 to 16-year-old pupils and ex-pupils, their parents, their teachers (and older leavers) were selected from these schools so that the probability of selection was equal for all members of each population," (Moreton-Williams and Finch, 1968, p. 247). Actual and theoretical distributions were shown to match closely and directly against types of school and geographical region. The social class distribution of the parents in this sample also coincided with that of the population at large. Interview success rates were high : pupils (96 per cent), teachers (99 per cent), parents (94 per cent).

Materials.

The interviews of teachers, parents, and pupils ranged across attitudes to school, perception of school objectives, the value and interest of subjects studied, secondary school curriculum, relations between home and school, careers advice and problems associated with raising the school leaving age with special attention focused on 15-year-old leavers. Data about demography and academic achievements were included.

Boredom was the focus of a special questionnaire. For 17 school subjects, pupils were first asked whether or not they were studying each. Then, of those being studied, they picked out which they thought useful, useless, or neither. Finally, they picked out the subjects they thought boring, interesting or neither (pp. 260-261, Q. 7).

A *Bore Score* was calculated as the number of subjects labelled boring expressed as a percentage of all subjects studied. For the main analysis a four-stage division was employed : 0–10 per cent, 11–20 per cent, 21–30 per cent, 31 per cent and above. The *Bore Score* was cross-tabulated against the variables whose detailed listing is precluded by limitations on space. Only the socio-economic grouping is specified.

Father's socio-economic grouping (Q.XX) (H3) : 7 value : The 1971 Census used 17 categories for locating jobs which were banded together in this analysis into 7 bands, the original Census numbers being in parentheses.

1. (1, 3, 4, 16) Employers and managers in large establishments, and all professional workers ;
2. (2, 5, 13) Employers and managers in small establishments, including farmer employers and managers, and intermediate non-manual workers ;
3. (6) Junior non-manual workers ;
4. (8, 12, 14) Foremen and supervisors, manual occupations and other workers on own account, including farmers ;
5. (9, 16 other than professional) Skilled manual workers ;
6. (7, 10, 15) Semi-skilled manual, personal service and agricultural workers ;
7. (11) Unskilled manual workers ;
8. (17) Unclassifiable ;
9. No information.

The selection of items was based on a mixture of criteria. Academic interest required that we look at certain sociological aspects of boredom and at school achievements of the pupils. However, administrative convenience, speed and cost also affected decisions. In the end, only three of the six possible data cards for each pupil were used such that they contained information from the three sources of pupil, parent and teacher and the estimates of boredom, achievement and social class. Other variables of interest on these cards were included.

Treatment of Results.

To comply (happily) with government regulations about the necessity of rendering individuals in social surveys unidentifiable, some variables were strongly banded before the data were transferred. The recombination of the 17 occupational categories into seven precluded some finer comparisons that might have been made, for example, rural/urban.

Computer programmes were written to calculate the *Bore Score* which was subsequently used to allocate pupils to four groups. A second programme generated cross-breaks of all other variables against this. After the elimination of 'no information' 'no answers' and 'don't know' and the occasional combination of rare values of other items, χ^2 analyses were run and supplemented with partitioning where necessary (see Maxwell, 1961).

Results are not described in detail and the only contrasts generally given are between the extremes, a percentage of the highest *Bore Score* (referred to, for simplicity, as *bored pupils*) is contrasted with that from the lowest (contrasting group). Many of the χ^2 values are way beyond the one in a 1,000 chance occurrence and some of these are complemented with the value of C, the Contingency Coefficient. Full details are to be found in an unpublished technical report (Robinson, 1974).

RESULTS

The original analyses of Moreton-Williams and Finch showed associations between the *Bore Score* and early leaving that made good sense, thus helping to establish the validity, and *ipso facto*, the reliability of the measure. Here we found that while only 45 per cent of *bored pupils* thought there were lots of interesting things going on in school, 74 per cent of the other pupils felt there were ($\chi^2=195$, df=3, $P<\cdot001$, $C=\cdot20$).

Tests of Hypotheses.

Each hypothesis was tested using one or more measures by calculating the value and significance of chi-square. As the vast majority of the tests showed statistical significance at beyond the ·01 level, details have been omitted from the text. They are shown in Table 1 in terms of an appropriate code number for the hypotheses. Non-significant values of χ^2 are shown in the text.

Immediate Antecedents. The single test of hypothesis A1 showed that 66 per cent of *bored pupils* claimed that school was the same day after day compared with 39 per cent of the contrasting group. The value of the contingency coefficient was 0·19.

Duffy's experimental studies (1972) showed associations between reported interest and value. Not only did adolescents claim greater interest in what they thought was valuable, they asked more questions and learned and remembered more about such matters. Here, it can be asked whether or not there are associations between usefulness and interest. Moreton-Williams and Finch point out aspects of independence between the two (p. 60), subjects can fail to be boring because they are enjoyable *per se*. While recognising the validity of a dissociation in both directions, certain disciplines may well be both useful and boring as well. Duffy's investigations require us to expect positive association overall—in so far as secondary schools are concerned to educate, we must argue that perceived uselessness should lead to boredom. The data were strongly consistent with this. *Bore Score* related very closely to the Uselessness Index (contingency coefficient=0·41 (A2a).

This was true at the same significance level for all partitions, higher Uselessness Indices going with higher *Bore Scores*, with the strongest contrast being between the highest *Bore Score*/Uselessness Index and the rest. Of those pupils who saw less than 10 per cent of subjects studied as useless, only 7 per cent were in the contrasting group, whereas 43 per cent of those seeing 30 per cent or more as useless were categorised as *bored pupils*. Further, 24 per cent of *bored pupils* made a general judgment that what was taught at school was useless compared with 7 per cent in low group (A2b).

TABLE 1

Chi-Square Values for Hypotheses Tested.

Hypothesis	χ^2	df	P
A1	166·0	3	·001
A2a	946·0	9	·001
A2b	138·9	3	·001
B1	11·4	6	NS (·1)
B2	21·6	6	·005
B3a	96·3	12	·001
B3b	85·9	3	·001
B3c	29·5	6	·001
B4	95·8	27	·001
B5a	77·6	15	·001
B5b	26·0	12	·025
B6a	121·3	3	·001
B6b	120·2	3	·001
B6c	187·4	3	·001
B6d	16·8	3	·001
B7	56·4	—	·001
B8a	43·1	3	·001
B8b	115·1	6	·001
B9	17·3	3	·001
Ca	10·2	3	·025
Cb	69·9	9	·001
T1a	120·2	3	·001
T1b	124·7	3	·001
T2	—	—	—
P1a	68·6	6	·001
P1b	16·0	9	NS (·1)
P2a	12·7	3	·01
P2b	9·4	3	·025
P2c	41·9	6	·001
P3	8·3	3	·05
Ha*	29·1	18	·05
Hb*	11·8	6	NS (·1)
Hc*	28·2	9	·001
Hd*	22·8	9	·01
He*	13·6	6	·05
Hf*	16·9	6	·01
H1a	15·4	6	·025
H1b	21·1	12	·05
H2a	26·0	21	NS
H2b	31·8	21	NS (·1)
H3	50·0	18	·001

* Not included as hypotheses in the model, although the variables are related to social class membership and home amenities (e.g., family size and parental education).

Behavioural Responses. The model shows several reactions to boredom with school subjects, all of which are supported in the data. The vast majority of pupils were not seen by teachers as either aggressive or withdrawn. Two per cent of pupils in the contrasting group were seen as very withdrawn or shy, 3 per cent of *bored pupils* were (B1) ($P < \cdot 10$). Nine per cent of the contrasting group were judged somewhat or very aggressive, rising to 16 per cent among *bored pupils* (B2). Other ' withdrawal ' measures showed stronger associations. Fifteen per cent of the *bored pupils* truanted occasionally or frequently, only 6 per cent of the others (B7). Twenty-three as opposed to 13 per cent of those allowed to do so had already left school (B8a), 59 per cent declared an intention to leave at 15 compared with 40 per cent in the contrasting group (B8b).

Bored pupils had a lower incidence of hard and very hard workers (32 *versus* 48 per cent) (B3a). The only index of aspiration was the taking of examinations. CSE did not discriminate, but whereas 73 per cent of the *bored pupils* were not taking any O-levels, this was true of only 58 per cent of the contrasting group ; 29 per cent of the latter were taking five or more O-levels, only 16 per cent of the former (B4). The teachers thought the most bored least capable : only 34 per cent of the *bored pupils* being judged as capable of at least two O-levels as compared with the 50 per cent in the other group (B5a). Intelligence test scores from ' Eleven Plus ' were available for only 53 per cent of the sample. Within this subsample, 34 per cent of the *bored pupils* had made scores of less than 95 compared with 25 per cent of the contrasting group (B5b).

Fifty-three per cent of the parents of *bored pupils* group claimed that their offspring did a lot of schoolwork at home, but this rose to 71 per cent in the other group (B3b). Forty-two per cent of the pupils with the highest *Bore Scores* read nothing, according to their parents, contrasting with 31 per cent of the lowest scores (B3c).

Bored pupils were generally more hostile to school than the other groups. They did not look forward most days to going to school (B6a) ; they got fed up more often with teachers telling you what you can and can't do (B6b) ; more were delighted when they had an opportunity to take a day off school (B6c); and most of their friends were more likely to come from near home rather than from the school (B6d). The differences between the bored and the not bored were substantial in size as well as statistically significant ; there appeared to be a general disenchantment with school as suggested by hypotheses B6.

Alas, this seems to extend to spare time as well. Although the 26 per cent of the *bored pupils* who admitted often being bored in their spare time, was not massively higher than the 20 per cent of the contrasting group, the difference is still statistically significant (B9).

Competition. With only the data to hand we cannot examine the extent to which boredom is linked to being a failure in a competitive system. We can note that 79 per cent of *bored pupils* were in streamed classes compared with 74 per cent of the contrasting group (Ca). 71 per cent were in classes where fewer than half the pupils were taking O-levels compared with 57 per cent of the contrasting group (Cb). Both results were more consistent with the proposed idea than its reverse. Neither age nor sex was associated with the *Bore Score.*

Teachers and School. Bored pupils were less likely to think their teachers took a great deal of interest in them (T1a) and more of them thought that teachers forgot they were growing up (T1b). Whether such perceptions were veridical we cannot say. We had no independent estimates of teacher interest and no information about teachers' personal boredom with the subjects they are teaching (T2).

Parents. Thirty-five per cent of the parents of the *bored pupils* were reported as showing no interest in their children's education, compared with only 20 per cent in the contrasting group (P1a) ; neither parent of 51 per cent of the *bored pupils* had discussed their child's education compared with 45 per cent in the other group. (P1b) (P<·10).

Parents were not asked directly about their concern about the education of their children. We have already seen that parents of the *bored pupils* report a lower incidence in their children of reading and of schoolwork being done at home ; it can be added that they did not think more time should be spent on these activities. And yet 90 per cent of parents whose children had high *Bore Scores* claimed that reading is important, which is only slightly lower than the 93 per cent of parents of pupils in the contrasting group (P3). However, neither fathers (P2a) (61 per cent *versus* 68 per cent), nor mothers (P2b) (66 per cent *versus* 70 per cent) of the *bored pupils* were as likely to claim that they themselves liked reading (P2). Their homes were less likely to have any books in them (P2c) (21 per cent *versus* 13 per cent).

On these weak indices, parents who had the most bored children were showing less interest in their education (P1) and were not so likely to pass on favourable attitudes towards education for their children to follow (P2).

Home. Very few of the characteristics of the household were associated with the *Bore Score.* Thirteen per cent of the *bored pupils* were from families with more than five compared with 8 per cent from the lowest group (Ha). Over 90 per cent of households had both parents present, and there was only a weak tendency for high *Bore Scores* to be associated with households where only the father was present (Hb) ($P < \cdot 10$). The position of the pupil in the family, whether parents were employed full-time, part-time, or not at all, where parents were married, divorced, widowed, or single, age of parents, were severally irrelevant to the *Bore Score.*

Parental ownership of the home was associated with a low *Bore Score* (H1a), but the number of bedrooms it had was not. However, having an excess of bedrooms over what was required was associated with low *Bore Scores* (H1b), as was living in a detached house (H1c). Since these differences are best expressed in terms of ' excess ' going with an absence of boredom, it follows that within the limitations of this sample, poor material conditions were *not* associated with high boredom with school subjects (H1).

Most of the indices of the amount and type of education of parents and siblings gave relationships with the *Bore Score*, but they were not, of course, independent of each other or of social class. Whether they function as modelling standards or as indices of present values or life style cannot be judged. Age at which father and mother completed school were more closely linked to *Bore Score* than were their qualifications or further education. Leaving at the statutory minimum provided the best single discriminating cut-off, 81 per cent of the fathers of *bored* pupils left at 14 compared with 74 per cent in the contrasting group (Hc) ; the figures for mother being 81 per cent and 73 per cent (Hd). Type of school attended was more weakly associated, both grammar and private schools reducing the probability of the child being bored (He—fathers ; Hf—mothers). Maternal further education and qualifications were not related to *Bore Score* at all.

Combined income of parents did not relate to *Bore Score* (H2a) and fathers' income only weakly so ($P < \cdot 10$) (H2b), and that association was in the middle range. Very high income provided no immunity against, and very low income no propensity for, high boredom.

Social Class. Both indices of fathers' occupation related to *Bore Score*, the seven-category split providing marginally more information than the traditional classification of the Registrar-General. The significance overall (H3) can be partitioned to show where the main significant differences lay (Table 2). There were no differences among classes I, II and III, the white collar jobs. A higher proportion of Class IV pupils had *Bore Scores* between 21 and 30 per cent, while Class V and particular Classes VI and VII were over-represented among *bored pupils*. Table 2 shows that it is the association between the highest *Bore Score* and having parents in semi-skilled and unskilled occupations that makes the greatest contribution to the relationship.

TABLE 2

PARTITION OF CHI-SQUARE FOR SOCIO-ECONOMIC GROUPING AGAINST BOREDOM INDEX (SIGNIFICANT APERTITIONS ONLY).

Boredom Group	Socio-Economic Groups	χ^2	P<
12 v 3	123 v 4	4·99	·05
123 v 4	1234 v 5	4·91	·05
123 v 4	12345 v 6	15·06	·001
123 v 4	123456 v 7	12·22	·001
Category 4 has highest *Bore Score*	Category 1 is the highest socio-economic group	(df=1)	

DISCUSSION

The results speak for themselves. That so many of the scores from the three separate sources of teacher, parent and pupil relate so meaningfully to the *Bore Score* attests to the reliability and validity of this index. We can, furthermore, feel confident that the examined concomitants, antecedents and consequences of boredom with school subjects offered in the model are a reasonable representation of the dynamics proposed.

The most important general feature of the model is its emphasis on the interdependence of the factors. While it may be helpful to ask about causes of boredom, it must be remembered that such factors may subsequently be affected by the boredom they created—giving the system a dynamic interdependence. Within itself it is a positive feedback model ; the cycle presumably reaches an unstable equilibrium only because external forces act to contain it. Escape for the bored pupil below the statutory school leaving age is not possible—the law and its sanctions presumably act eventually as a major force to keep the bored pupil ticking over in the classroom.

Several points of minor significance may be noted. We had no information about the physical amenities of school, but might expect them to be of marginal relevance only in this regard. As in the work of Moreton-Williams (1967), Wiseman (1964) and others, we found the material and financial conditions of the home to be of little import. It was the values, interests and behaviour of parents that were.

It is unfortunate that there are no independent estimates of these same characteristics in teachers, only what the pupils see. Why should they misperceive ? Why should they lie ? No doubt there are reasons, but it is perhaps time to begin to ask how many teachers are either not concerned with teaching, are uninterested in the subjects they are teaching, or do not have any faith in the learning possibilities of their pupils. To do so is not to locate blame or responsibility, but to examine the situation with a view to suggesting possible remedies.

The relevance of parents and teachers as models for pupils should not be underestimated. The popular stereotype of rebellious youth despising authority in any form does not correspond to the facts. As Spencer (1972) showed in his study of attitudes of adolescents, the majority aspire to orthodox ways of leading satisfying lives and are content to endorse the values of parents and schools.

The association between boredom and competition is also noteworthy. Where competition between individuals is constructed and encouraged and where winning is rewarded, it is inevitable that there must be losers. But is it not absurd to construct a system of education where only a minority are viewed as winners ? And what are the chronic losers to do ? Very few people will persist in playing at some game in which they invariably lose, and yet we tolerate a situation in which children are consistently allowed to lose every school day, for perhaps ten years of their lives. That some children become very bored is not surprising ; that their behaviour is generally so restrained perhaps is. What little evidence we have is consistent with the idea that inter-individual competition increases the likelihood of some children becoming bored.

While intelligence test scores at ' Eleven Plus ' did relate to *Bore Score*, the association was much weaker than those between *Bore Scores* and present and predicted attainments. High IQ Scores seemed to provide some immunity from boredom, but very low IQ (below 84) was not associated with higher boredom. If we were to concede that IQ scores were a measure of potential rather than of interaction between potential and opportunity taken, we could still argue that ability *per se* is not a very important predictor of boredom. Being bright helps to save, but being dull does not damn.

And who is likely to say they are bored by a substantial amount of what they study in school ? We should expect a relatively high incidence of such replies from lower working class pupils whose parents show little interest in education in general or in the education of their own children in particular, and who themselves left school at the first opportunity. We should expect to find them in the lower streams of streamed schools : possibly schools where the emphasis is on O-level achievement and the lower streams do not take such examinations.

Such considerations as these seem to thrust any responsibility for initiating change on to adults rather than pupils—in the first instance at least. If we assume that pupils would rather not be bored—and to argue the contrary is probably a symptom of defensive mental acrobatics—we can see that changes need to occur. We have the power to reduce inter-individual competition and the chronic failure inherent to such a system. Parents and teachers could show more interest in the education of their charges.

We have a means of diagnosing boredom. We can simply ask pupils whether or not they are bored. We can prevent boredom by showing pupils that what is being taught is valuable and useful—and by changing the structure and content if it is not.

As psychologists, we may note the potential of boredom as a field of study to excite curiosity.

ACKNOWLEDGMENTS—This work was supported by a grant from the Schools Council, London, to whom grateful acknowledgment is made. The analyses of the data relied on computer programmes developed through the good offices of Dr. C. D. Creed and Mrs. A. Wardle.

REFERENCES

Atkinson, J. W. (1964). *Introduction to Motivation.* New York : Van Nostrand.

Bandura, A., and Walters, R. H. (1963). *Social Learning and Personality Development.* New York : Holt.

Berlyne, D. E. (1960). *Conflict, Arousal and Curiosity.* New York : McGraw-Hill.

Crowther Report (1959). *15 to 18.* London : HMSO.

Hadow Report (1931). *The Primary School.* London : HMSO.

Duffy, M. G. (1972). Expression of curiosity, interest and retention. Unpublished masters' thesis. University of Southampton.

Maxwell, A. E. (1961). *Analysing Qualitative Data.* London : Methuen.

Morton-Williams, R. (1967). The 1964 National Survey : Survey among parents of primary school children. In the Plowden Report, *Children and their Primary Schools, Vol.* 2. London : HMSO.

Moreton-Williams, R., and Finch, S. (1968). *Young School Leavers.* London : HMSO.

Newsom Report (1963). *Half our Future.* London : HMSO.

Plowden Report (1967). *Children and their Primary Schools.* London : HMSO.

Robinson, W. P. (Ed.) (1974). Education, curiosity and questioning. (Unpublished technical report.)

Spencer, C. P. (1972). Selective secondary education, social class and the development of adolescent sub-cultures. *Br. J. educ. Psychol.*, 42, 1-13.

Wiseman, S. (1964). *Education and Environment.* Manchester : Manchester University Press.

(*Manuscript received 20th June,* 1974)

Br. J. educ. Psychol., 45, 153-161, 1975

CONFIRMATION OF A GENERAL FACTOR OF EFFECTIVENESS-MOTIVATION BY INDIVIDUAL TESTS

By D. H. STOTT
(*Department of Psychology, University of Guelph, Ontario*)

AND J. B. ALBIN
(*Department of Psychology, University of Western Ontario*)

SUMMARY. Following the finding of a general factor in the 11 areas of function covered by the Stott-Sharp Effectiveness-Motivation Scale for pre-school children, five individual tests exploring various aspects of effectiveness motivation were devised. These were applied to two samples of 4-year-old children (N=82). All five tests correlated positively with the E–M Scale, and all except one test correlated with the others at a generally satisfactory level of significance. These results were held to confirm the existence of a motivational factor which has been named effectiveness.

INTRODUCTION

THE concept of a general motivation towards competence or effectiveness was first put forward by White (1959). He defined it as, " an intensive need to deal with the environment . . . (The child) selects for continuous treatment those aspects of his environment which he finds it possible to affect in some way." Later in the same article he speaks of " the selective continuance of transactions in which the animal or child has a more or less pronounced effect upon the environment." While White's formulation aroused considerable interest, it did not gain the recognition it deserved as a more comprehensive explanation than the segmental explanations heretofore advanced (aggression, curiosity, stimulation, manipulation, etc.). One reason for this was that, although his concept was empirically induced, he did not translate it into systematic observations that could be the subject of experimental verification.

Soon after White's paper, Stott (1961) published his account of his son's ' non-organic ' behaviour from near birth until about the age of 17 months, from which he induced a very similar concept, naming it effectiveness-motivation. His method was to record all behaviour which seemed motivated, yet did not satisfy any visceral or social need. Evidence that the behaviour was motivated, rather than random ' exercise,' was obtained by the child's reaction to its outcome in the form of perseverance or unambiguous repetition, expressions of gratification or of annoyance at failure, and precedence over organic needs. He grouped the observed behaviours by 20 classes of outcome, and then reduced these to five broad categories of non-organic motivation. He then claimed that these were " aspects of a general *relationship of effectiveness* which the child seeks to maintain or establish between itself and its environment."

Stott and Sharp (1968) compiled the *Effectiveness Motivation Scale*, which is a check-list of the spontaneous behaviour of pre-school children in 11 areas of interreaction, scaled to cover varying manifestations of effectiveness. Stott, Williams and Sharp (1974) found that the area scores all correlated significantly with the total effectiveness score less the particular area score. They interpreted this as consistent with the concept of a general motivation towards establishing effective relationships with the environment.

The observations on which this scale was based were those of the pre-school teachers. Consequently, the alternative explanation, that the correlations were due to a halo effect, even though unlikely, could not be dismissed, and the hypothesis that the general factor was a motivation towards effectiveness required independent confirmation under controlled test conditions.

AIM AND METHOD

The aim of the present study was to confirm or otherwise the presence in the young child of a general motivational factor which might be described as a tendency to establish relationships of effectiveness with his world. This involved placing the child in various test situations which gave him the opportunity to demonstrate effectiveness in externally dissimilar types of transaction for which no visceral or social motivation seemed likely. If the performance of children on a series of such tests intercorrelated, this would be consistent with a general factor of effectiveness motivation.

Stott and Albin designed five pieces of apparatus which would give the pre-school child opportunities to establish one or more of the five categories of effective relationship proposed by Stott (1961). These were :

1. *Recognition and discrimination*, showing some pleasure or contentment at recognising an object.
2. *Completion*, realisation of some envisaged arrangement such as construction, fitting or completing a task.
3. *Control*, exercising physical power or control over the inanimate environment.
4. *Exploration*, resulting in gains in knowledge about the environment as well as manipulatory and curiosity behaviours.
5. *Effecting change*, making a loud noise, reducing an object to fragments or doing anything which leaves environment different.

The third category was later extended to include the child's control over his own impulse system, notably fear, documented by Murphy (1962, pp. 168-190). Other criteria for the selection of the tasks were (*a*) that in order to minimise previous learning opportunities they would have to be novel ; and (*b*) they should appeal to the subjects to the extent that they would be eager or at least willing to undertake them.

Test Apparatus.

The apparatus and their concomitant tasks were improved during the testing of an initial sample of subjects and took their final form in the following tasks :

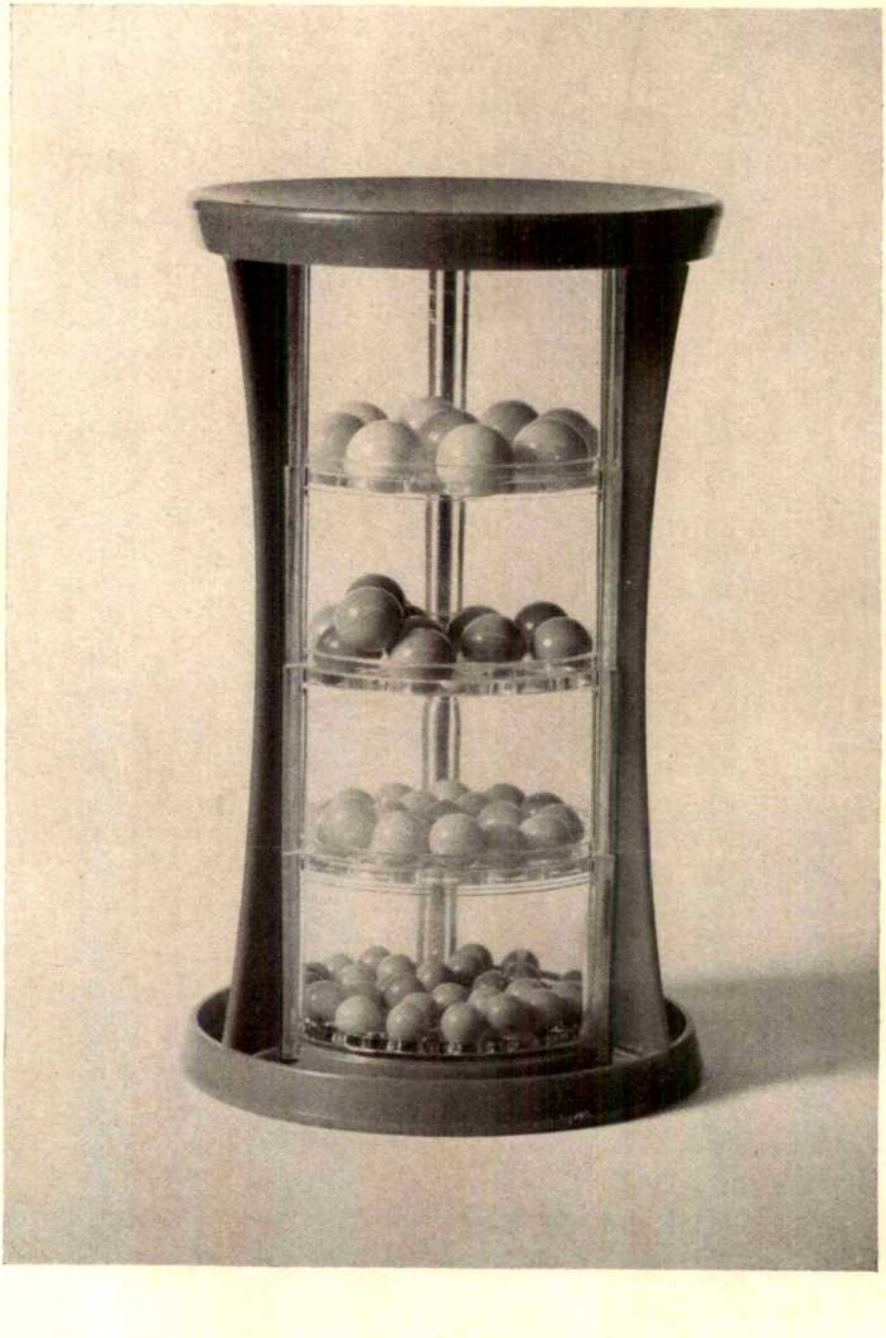
(1) *The Noise Box.*

This consisted of a closed wooden box $15\frac{1}{2}\times10\times5$ inches which produced some resonance. On its upper surface, clearly visible and accessible, were the following noise-producing attachments : a pair of spring-steel shoe trees which oscillated and clapped together when manipulated ; an electrically operated bell button ; four guitar-type strings ; a buzzer button which emitted low-frequency sound vibrations ; a button operating a standard buzzer ; and a pair of wire-spring door stops which vibrated when bent and released.

After the box was placed in front of the subject he was told that it made many different noises, and he was invited to play with it. The subject's performance was scored as follows, the type of effectiveness-behaviour evoked being given in brackets :

FIGURE 1

THE MARBLE SORTER



(i) *Willingness to start manipulating the components (producing a noteworthy effect)* :

If the child needed verbal encouragement to begin	0 points
If the child asked questions as if to make sure he was allowed to begin	2 points
If the child started without hestitation on being given the instructions	4 points

(ii) *Number of components manipulated (exploration)* .... 0–6 points

(iii) *Patterning of effects (realisation of an envisaged arrangement) and producing different effects (recognition and discrimination)* :

For a simple continued effect with one component	2 points
For a complex effect (tune or beat) with one component	4 points
For a complex effect (tune or beat) with two components	6 points
For a complex effect (tune or beat) with three components	8 points

(iv) *Mastery of fear of the apparatus (control of own impulse-system)*:

If child needed verbal encouragement to continue	0 points
If child constantly looked to the experimenter as if for permission to continue	2 points
If, once the child had started, no encouragement was needed	4 points

It is seen that this piece of apparatus covered in some measure all five of Stott's categories of effectiveness relationship.

(2) *Map Shapes.*

Twenty fragments of maps in outline were presented to the child in succession. He was told that the shapes might have been drawn by children, and asked to help identify them. Some of the shapes were strongly suggestive of objects, others not suggestive at all. For each distinct object named (dog, cow, boot, etc.) three points were given, two for the first repetition of an object and one for subsequent repetitions. The naming of the 20 subjects all differently would give a maximum of 60 points, but nowhere near this was ever achieved.

It was judged that this task would test the strength of a child's motivation to recognise and discriminate, with an element of make-believe or imagination (representing a strong motivation to recognise) in cases where the shape did not obviously suggest any object.

(3) *Marble Sorter.*

This is marketed as a ' Colour Sorter ' by Child Guidance Products Inc. of Bronx, New York (Figure 1). It is a transparent cylinder 6½ inches high, closed at each end, and divided internally by three floors. Within each of the floors are holes of varying size which permit coloured balls of smaller sizes to pass through. There is in addition a base floor with seatings for the balls but no holes. The balls which are too large are retained in successive compartments, which amounts to a sorting by four colours.

The marble sorter having been placed in front of the child, he was told that it was a magic toy and given permission to play with or do anything he liked with it as long as he kept it in his hands. This he was allowed to do until he stopped or indicated that he had finished.

Scoring depended upon the completeness of the sorting produced or aimed at by the subject, as follows :

(i)	Shaking sorter in a way that produced noise only, without attention to any sorting effect	1 point
(ii)	Reversing sorter and watching with some concentration as the marbles trickled down, but with no evident effort to produce a sorting	2 points
(iii)	Turning the sorter quickly back to its original position after reversing so that the marbles return to the unsorted chamber	3 points
(iv)	Holding sorter in reversed position so that the marbles are on the whole distributed by floors, but without seeming to try to produce a complete sorting	4 points
(v)	Shaking the sorter in a way that produces a fairly good sorting, but not achieving a perfect one	4 points
(vi)	Shaking the sorter until a perfect distribution by floors is effected	5 points
(vii)	Once the marbles had been sorted by floors, attempts, even though soon abandoned, to seat them in the holes through which they are too big to pass	6 points

It was judged that the marble sorter task would cover the following categories of effectiveness-motivation :

Producing a spectacular effect, visually and auditorially.

Completing an envisaged schema.

Discrimination of colours and recognition of a pattern.

(4) *Matchers Game.*

The Matchers game is an item of the Flying Start Learning-to-Learn Kit (Stott, 1972) designed to teach reflectivity. It consisted of duplicate sets of six cards, 2¼ × 4½ inches, each of which contained a variation, for example, a pirate with and without a wooden leg. The variations were not such as to approach the limits of the child's perceptual discrimination, but in order to choose the right card three criteria had to be taken into account. One set was placed on the floor before the child. The experimenter presented one of the duplicate cards at a time and asked the child to find the card that was exactly the same from among those laid out on the floor. The child checked the correctness of his choice by the number of bars on the backs of the cards. The experimenter similarly laid out a set of six cards for himself, and the child handed him the cue cards for matching. Cards matched correctly were retained by the player, those incorrectly matched being handed back.

The scoring system took no account of the child's success, but only of his willingness to participate. It is summarised below. If the child, in choosing a card :

Is exceedingly apprehensive, and needs a lot of encouragement, looks at experimenter for assurance prior to making a choice, etc.	1 point
Needs encouragement to make choice	2 points
Hesitates but finally chooses without being prompted.	3 points
Hesitates only long enough to be sure of being correct	4 points

Following the compensatory principle used in the scoring of the Effectiveness-Motivation Scale, where the child's motivation or problem-solving style was such as to render the above scoring inapplicable (e.g., when the child made a quick guess without being concerned with success or failure, or was unwilling for other reasons than apprehensiveness) a compensatory score of 2 was given. Further points, 1 to 4, were given for the acts of giving the experimenter his cue card, on analogous criteria.

The above scoring system registered the child's willingness (1) to participate in a game where he could succeed or fail, win or lose, that is to say, to make a bid for control and dominance, and (2) to make recognitions and discriminations.

(5) *The Form Boards Persistence Tasks.*

A simple and a more complex Form Board were used. The first required a near-square of wood, 6 × 6¼ × ½ inch, to be fitted into the corresponding shape, but turned through an angle of 90°. Since the slight deviation from a square was not visually discriminable to the children, their first trial was unsuccessful in all cases. The more complex task consisted in fitting 12 pieces of wood into a near-square 7½ × 8 inches. All the pieces were 1 inch wide but varied in length from 1½ to 6 inches. The scoring measured the subject's willingness to make the attempt and persistence in completing the puzzles, irrespective of actual success or failure, on two scales from 0 to 3 for each form board.

The Effectiveness-Motivation Scale.

Immediately after the individual testing was concluded, the children's teachers completed an Effectiveness-Motivation Scale in respect of the play activities of each child as observed over the previous month or so. Following the standard method of scoring, the subject was given from 0 to 4 points in each of the 11 areas of function, amounting to a maximum effectiveness score of 44.

The Samples.

The above tests were administered initially to 21 children (11 males and 10 females) aged 4 to 4½ years who had attended a pre-school for at least three months. All their parents belonged to the middle or upper-middle class. After certain minor improvements had been made in the procedures and scoring, they were re-administered to 61 children aged between 4 years and 4 years 9 months, 30 of whom (21 males and 9 females) attended a day care nursery, and 31 (16 males and 15 females) attended a co-operative pre-school. Their socio-economic status was also middle and upper-middle class. Nearly all the children in both samples were Caucasian and all were English-speaking.

RESULTS

Pearson product-moment correlation co-efficients were used to correlate the teachers' assessments made on the Effectiveness-Motivation Scale with the children's scores on the five individual tests. These correlations are given in Table 1. All are statistically significant, barring that involving the Map Shapes with the first sample ; those involving the Noise Box and the Persistence Form Boards are highly significant.

TABLE 1

Correlations of Five Individual Tests with Effectiveness-Motivation Scale.

Apparatus	Correlation with E-M Scale	
	First sample N=21	Second sample N=61
Noise Box	0·72**	0·74**
Marble Sorter	0·40	0·37**
Map Shapes	0·16	0·31*
Matchers Game	0·69**	0·38**
Persistence Form Boards	0·60*	0·63**
All five above	0·78**	0·74**

Significance : P<*·05 **·01

The intercorrelations between the five individual tests are given in Table 2. Those for the first sample are given in parentheses. The Noise Box test once again shows the highest correlations with the other postulated measures of effectiveness motivation, and the Map Shapes the least. Except when this test is involved, no inter-test correlation falls below the ·01 level of significance.

TABLE 2

Intercorrelations Between Pairs of Individual Tests. (N=61 ; 21)

Apparatus	Marble Sorter	Map Shapes	Matchers Game	Persistence Form Boards
Noise Box	0·49** (0·35)	0·48** (0·06)	0·48** (0·5*)	0·68** (0·27)
Marble Sorter		0·22 (0·36)	0·32** (0·15)	0·47** (0·22)
Map Shapes			0·13 (0·31)	0·30* (0·19)
Matchers Game				0·49** (0·42)

Significance : P<*·05 **·01
Correlations for the first sample (N=21) are given in brackets.

DISCUSSION

From the observation of eleven areas of children's spontaneous behaviour Stott, Williams and Sharp (1974) demonstrated a general factor which they termed effectiveness motivation. The purpose of the present study was to confirm or disconfirm the existence of such a general factor.

These experimental tests showed a high degree of intercorrelation with the children's spontaneous behaviour as recorded on the Effectiveness-Motivation Scale. This was consistent with the hypothesis of a general factor. The correlations were on the whole higher than usually obtained in the testing of pre-school children. However, before a claim can be made that they demonstrated individual differences in the strength of effectiveness seeking, other possible reasons for the relationships must be examined. Correlations may be inflated by similarities between methods of testing, rather than from a general factor mediating the subjects' responses. Campbell and Fiske (1959) suggest, for example, that if a series of questionnaires is used, significant correlations may result from the subjects' reactions to questionnaires themselves rather than from relationships between the variables being measured. As will be shown in detail below, the present series of tests presented the child with reasonably diverse opportunities for exercising effectiveness. In the Noise Box it was demonstrated through the auditory modality, while in the Marble Sorter it was primarily visual. The fitting of pieces required on the Form Boards would probably be verified tactually. Even when the visual modality was involved, the types of effect were qualitatively diverse. With the Marble Sorter it was patterning, with the Matchers Game it was discrimination of differences, while with the Map Shapes the recognition or imagination of similarities was predominant. The correlations obtained could, therefore, hardly reflect a similarity of test or method.

A further consideration is whether the correlations represent a general factor of ' intelligence.' It will be noted, first, that the scoring took no account of the success of the child in any of the tasks. They were thus not measures of cognitive ability. Nevertheless, it is likely that effectiveness motivation will be found to correlate with IQ. Kronstadt (1973) indeed found a correlation of 0·25 ($P < \cdot 05$) between scores on the Effectiveness-Motivation Scale and the Binet Test in a sample of 63 disadvantaged black 3-year-olds. If we accept J. McV. Hunt's (1961) concept of intelligence as " comprising strategies for processing information that develop in the course of the child's interacting with his environment," it follows that a child's cognitive skills at any stage of development are a function of his motivational history. The extent to which his cognitive abilities have been brought into use, and hence developed, will have depended on his willingness to face novel and complex tasks and the degree to which he is reinforced by success or daunted by failure. It could thus be argued that the IQ is a measure of motivational strength as well as of cognitive efficiency or quality (Stott, 1971). If the IQ is seen as a means of recording the growth of mental competence, the issue of whether there is a factor of ' intelligence ' in effectiveness motivation, or of effectiveness motivation in 'intelligence,' becomes a false one, since the two factors may well be mutually dependent. In our exploration of the determinants of mental growth, we need a means of measuring the motivational variable which discounts so far as may be the child's attainment in specific tasks.

It has next to be asked to what extent the results were consistent with concept of competence or effectiveness motivation as induced by White and Stott. In Table 3 the types of reinforcing outcomes of each of the test tasks are summarised in terms of those postulated by Stott (1961) as the basic manifestations of effectiveness motivation. It is of interest that the Noise Box, which taps all five categories of effectiveness-motivation, correlated most highly both with the other four individual tests and with the Effectiveness-Motivation Scale.

The low correlations of the Map Shapes with the other tests raises the question of its value in an effectiveness motivation test battery. It is, however, the only test which involves an element of make-believe—that is to say, an imaginary reconstruction of the environment in a way which facilitates fulfilment of a need. Such a strategy may be used only by a minority of children, but it provides theoretically good evidence of strength of motivation.

The Marble Sorter is the only test which includes visual patterning as a form of closure or realisation of a schema. General observation suggests that only certain children, and indeed adults, are interested in discerning visual patterns, even though they may have artistic interests of other types. We do not yet know much about such preferences and channels of fulfilment, except that they exist to a marked degree even in young children. There are thus good reasons for retaining test items which sample diverse areas of fulfilment, even though they have only modest correlations with other items.

TABLE 3

CLASSES OF REINFORCING OUTCOME RELATED TO EACH TEST.

Apparatus	Reinforcing outcome
Noise Box	1. Recognition and discrimination in the auditory modality. 2. Completion (achieving an envisaged schema in the form of rhythm or tune). 3. Control (of own fears). 4. Exploration (of the effects of manipulation). 5. Effecting change (making spectacular noises).
Map Shapes	1. Recognition and discrimination, to the point of make-believe, in the visual modality.
Marble Sorter	1. Recognition and discrimination (of colour and pattern). 2. Completion of an envisaged schema (in the form of a visual pattern or arrangement). 3. Effecting change (of a spectacular auditory or visual character).
Matchers Game	1. Control (in the sense of accepting a challenge to win, exercise dominance). 2. Recognition and discrimination.
Persistence Form Boards	1. Completion.

We have finally to consider how the statistical findings can be translated into a psychologically or biologically feasible concept. The theory of effectiveness motivation implies that each individual, during his normal waking state, appraises his relationships to his environment by certain criteria—of recognising and discerning differences, finishing some task, controlling objects and people (including himself), in a way that he is making the running, knowing and understanding, and effecting worthwhile changes. All such transactions place him in a more advantageous position for achieving goals and for maintaining homeostasis of a more materially biological character. We have to suppose that such criteria for appraisal were evolved from quite simple sensory cues. These could have been the recognition of specific objects, completing particular cycles of activity once they had begun, dominance in a physical sense as seen in infra-human animals, being physically in a more elevated position than one's congeners, exploration of the creature's whereabouts with a view to specialised knowledge regarding safety and food supply, and merely seeing effects of physical intervention (as with the love of destroying things observed by Köhler in his apes). Owing to the large number of somewhat similar advantageous situations of each type, there would be an economy in generalising about them as desirable goals. To take just one example : " I'm King of the Castle " is played by many species of animals in various forms as presumably symbolising status ; in human adults an analogous gratification is afforded by constructing high towers or by being in an ' elevated ' social or economic position.

It is also reasonable to suppose that—for economy of neural structure—a common reinforcing mechanism would be evolved. Part of it would be to provide feelings of gratification which are the counterpart of feelings associated with physical comfort, taste or sexual pleasure. White has given the name of effectance to this feeling-aspect of competence or effectiveness.

How far such generalisation of effectiveness goes will be discovered only when we have been able to map out this area of motivation. As in all biological phenomena we shall probably find a state of affairs which does not fit our preference for simple conceptualisation. Effectiveness motivation may be a kind of conglomerate which has gone through a degree of fusion but retains the elements of specific types of material goal. In view of this possibility, its mapping out will require a widely ranging testing programme.

REFERENCES

CAMPBELL, D., and FISKE, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.*, 56, 81-105.

HUNT, J. McV. (1961). *Intelligence and Experience.* New York : Ronald Press.

KRONSTADT, D. C. (1973). The relationship of a motivation to achieve, intellectual functioning, language ability and behavioural functioning in 3-year-old disadvantaged children. Unpublished Ed.D. Dissertation, University of Florida College of Education.

MURPHY, L. B. (1962). *The Widening World of Childhood.* Chicago : Basic Books.

STOTT, D. H. (1961). An empirical approach to motivation based on the behaviour of a young child. *J. child Psychol. Psychiat.*, 2, 97-117.

STOTT, D. H. (1971). Behavioural aspects of learning disabilities : assessment and remediation. *Amer. Psychol. Assoc. Experimental Publication System*, 11, 400-36.

STOTT, D. H. (1972). *Flying Start Learning-to-learn Kit.* Guelph, Ontario : Brook Educational ; Edinburgh, Holmes McDougall.

STOTT, D. H., and SHARP, J. (1968). *Effectiveness-Motivation Scale.* (Publication pending).

STOTT, D. H., WILLIAMS, H. L., and SHARP, J. D. (1974). A general factor termed effectiveness motivation in the behaviour of pre-school children. *Educ. Res.*, in press.

WHITE, R. W. (1959). Motivation reconsidered : the concept of competence. *Psychol. Rev.*, 66, 297-333.

(*Manuscript received 25th July, 1974*)

Br. J. educ. Psychol., 45, 162-176, 1975

COGNITIVE PSYCHOLOGY AND EDUCATION

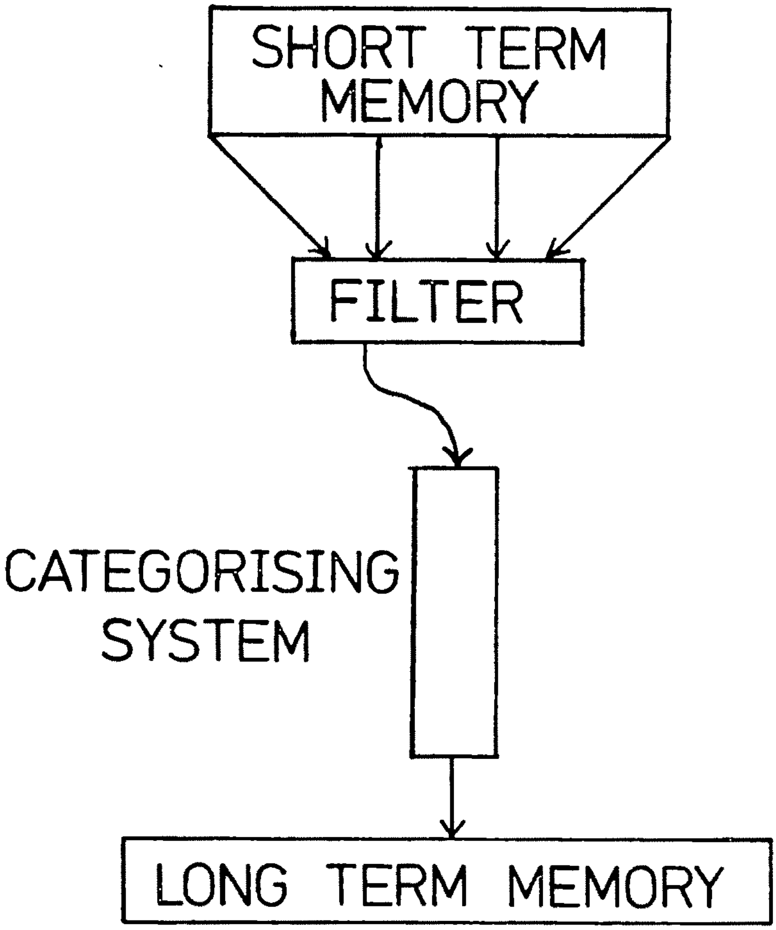
BY D. E. BROADBENT
(*Department of Experimental Psychology, University of Oxford*)*

INTRODUCTION

EACH of the areas mentioned in this title is a large one, and obviously we shall consider only some aspects of each. The broad field covered by ' Education ' is already familiar : it can stretch from the primary school to the university, with forays beyond the frontier in either direction, and it can take in an enormous range of attitudes amongst educators. One could caricature the position at one extreme as that of the teacher as behaviour modifier, manipulating each shade of the pupil's behaviour with appropriate schedules of reinforcement : while the opposite extreme, and perhaps the more popular nowadays, would be the teacher who does not so much instil knowledge as make it available. If his pupils should choose to want it they can find it. As may be suspected, one purpose of this paper is to suggest some rationale for rejecting these extremes.

FIGURE 1

A TYPICAL ' BLOCK DIAGRAM ' THEORY OF COGNITIVE PROCESSES (BASED ON BROADBENT, 1958).



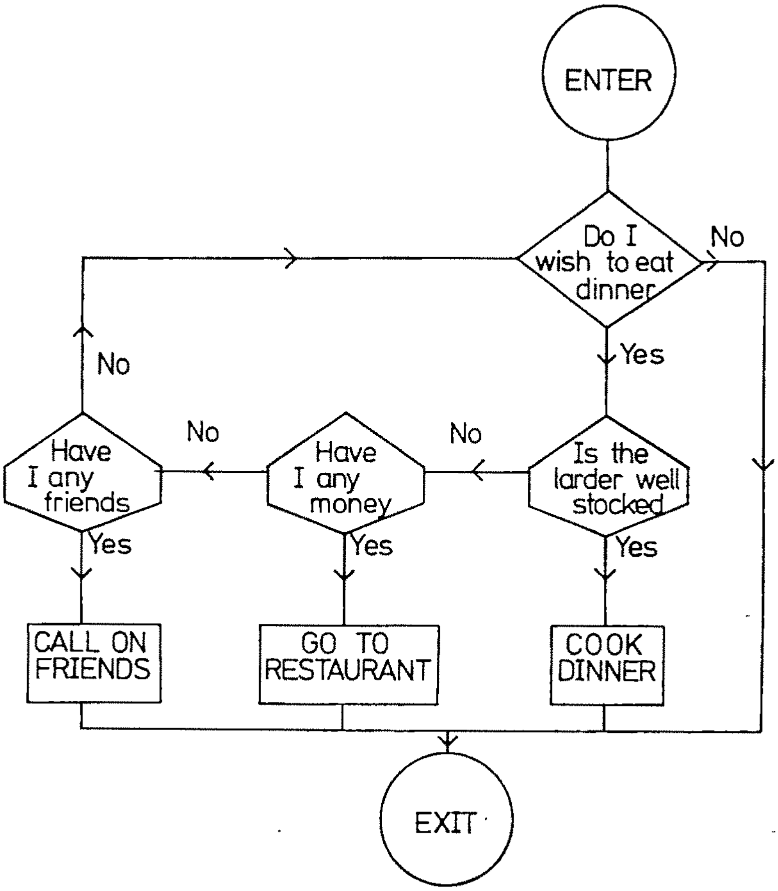
---

* The author is on the external staff of the Medical Research Council.

The rationale is to be found in 'Cognitive Psychology,' which is itself almost as vague a term as 'Education.' Taking our start from the classic textbook of Neisser (1967), we can say that it includes all those parts of experimental psychology which have to do with complex perception, the organisation of memory, language, thinking, and in general the complicated processes which underlie the intellectual side of human life. Investigators working in this area tend to express their theoretical findings in one of two ways. First, they may draw a *block diagram* consisting of a series of boxes, each representing a function or process going on within the nervous system, and any particular item of information is supposed to find its way from block to block as it is processed. A typical example is shown in Figure 1, which represents a simple outline of some of the views of Broadbent (1958, 1971). In this diagram, information is supposed first of all to enter a short-term storage system which will hold a number of items of information simultaneously : some of these items are selected from this short-term store by a filter, and pass through a system for categorising or encoding the item, so that the same central event will occur whether the man sees the word 'two' or '2' or 'II' or hears a male or female voice saying the word 'two.' From the categorising system, the information is supposed to go to long-term memory. Thus, a particular item of information moves from one place to another within the system.

FIGURE 2

A TYPICAL 'FLOW CHART' THEORY OF COGNITIVE PROCESSES.



The other kind of diagram drawn by cognitive psychologists is a *flow chart* like that in Figure 2. Here we are dealing no longer with information moving from place to place, but with events at one place at successive moments in time. In this example we see the successive events which might go on when a man is deciding on his future course of action at some time in the early evening. He might decide successively that he wished to eat his dinner, that he had no food in his larder, but that he did have money and would therefore go to a restaurant, or he might take a different decision at any point and go through a different set of actions. Indeed, he might get into an endless loop which can only be broken by his deciding that he does not wish to eat dinner after all ! Notice that rather different kinds of behaviour would be produced by a man who thought about imposing on his friends before he considered the state of his own larder, rather than the opposite sequence, and one could, therefore, examine the actual behaviour to see whether it agrees with the theoretical model. The diagram does, however, represent a series of events in the same place at successive times, rather than a series of places in which the same information can be located : it involves time rather than space. An example of a flow chart for a more complicated piece of human behaviour is given by Winograd (1972, p. 64) for the case of a computer analysing a sentence and deciding on the tense of the verb.

Both kinds of diagram are extremely useful in representing what happens inside people. However, they very rapidly become extremely complicated, and another purpose of the present paper is to point out another way in which psychological events can be described.

With these background points in mind let us consider three fields in which cognitive psychology can be related to education.

THE MICROSTRUCTURE OF TEACHING

If we think of a particular child in one school at one time, he or she is surrounded by a vast array of incoming signals, perhaps especially so in a modern open classroom. In general terms, however, Figure 1 represents a genuine limitation on his performance : he cannot process simultaneously everything that reaches his senses, and has to select from what is present. If something is not noticed at the time it happens, it has hardly any chance of affecting long-term memory (Moray, 1959 ; Turvey, 1967). Yet it is by no means automatic that the particular stimulus which interests the teacher will be the one which gets access to the processing centres of the child. Even an adult, when faced by strange visual patterns differing in a number of features, tends to classify them by the features which vary most conspicuously (Imai and Garner, 1965) ; and similarly it would be optimistic to expect the child to notice inconspicuous features of visually presented letters, when there are large bright objects also in the visual scene, or to listen to a quiet and monotonous voice explaining a problem, when there are loud and suddenly changing bursts of laughter or traffic noise occurring at intervals. The extreme view of education described earlier, in which information is made available for the enquiring mind to discover, can only be a caricature because the information has to be presented in some physical form. That form and the background of stimulation which surrounds it will, in fact, decide whether the child attends to this item of information or to something else in the surroundings. No matter how apparently undirective a school or a home may appear to be, biases of this kind exist, and should be handled deliberately rather than left to chance.

Indeed, the doubt which many of us would have about a traditional 'chalk-and-talk' classroom would be that it violates some of the principles of control of attention ; the use of low intensity stimulation continuing for a prolonged period from a single direction, the lack of link to any existing interest of the child (see Moray, 1959, again) and so on. The danger to be avoided is that, in changing such a situation, one might still fail to present information in such a way that is noticed. Audio-visual aids, for example, can indeed replace the traditional classroom by intense, changing and interesting stimuli, yet laboratory studies of attention would lead us to expect that there is still a limit to the amount of information which can be presented in a certain length of time, and merely by adding stimulation of the ears to that of the eyes we cannot expect twice as much to be assimilated. If what is being heard is truly independent of what is being seen, then we can take in either but not both (see for example Mowbray, 1953). The proper use of sound and vision together is rather to provide variety of stimulation by changing at intervals from one channel to the other, or else to present to eye and ear signals which are already closely related to each other in the child's experience, such as words being spoken by a face which is seen or words which pick out one part of the visual scene for special attention.

Most of the experiments which make these points have been performed with adults, yet the work of Maccoby and Konrad (1967) shows that similar principles apply with quite young children, and indeed, that younger children are less helped in attention by the occurrence of a sequence of probable events. As yet they do not know that a word dealing with a particular topic is likely to be followed by another word on that same topic, and the task of holding attention on a single problem may, therefore, be harder for them than uninstructed adults will realise. Experienced teachers already understand the importance of the physical facts underlying attention ; perhaps they too may forget that their inexperienced colleagues do not know these things ! The laboratory experiments of cognitive psychologists, therefore, can help us to design the detailed environment in which learning takes place, with a slightly better chance that the child will notice and react to the information that is being presented.

Even when a child has noticed something about a situation, the way it has been noticed is important. Morton (1967) has a light-hearted demonstration of the importance of this factor, in a study which involved asking regular telephone users to reproduce on the telephone dial the positions of the letters and digits. Although they have seen and reacted to these stimuli many times, most of them have not learned which holes in the dial contained which letters. They had, of course, no need to do so: their response to the letter when dialling it was merely to put a finger in the appropriate hole, and not to organise the information in some form appropriate to long-term memory. Here again, one can see the reasons for discontent with the traditional classroom. Suppose a child is presented with a list of the kings of England and merely processes this information in such a way as to pass a classroom test at the end of the week, it does not follow that his behaviour has been affected in any other situation. Yet, equally, one might deduce from Morton's experiments that some modern alternatives to drill are ineffective. A child who repeatedly discovers the results of multiplications by using rods or repeated addition will not therefore acquire a long-term store of information sufficient to check change in the supermarket.

There are many laboratory examples of a similar kind. For example, you can be made much less efficient at recognising that the word 'jam' has been presented to you on a previous occasion, if last time it appeared it was accompanied by the word 'traffic' while on the test it was accompanied by the word 'raspberry' (Light and Carter-Sobell, 1970). Cognitive psychology not only warns us that information we present may fail to get into the nervous system at all, but also makes it clear that the future effects of what is seen or heard depend very much on the categories or codes into which the information is placed. Although we started by looking at an individual child in a particular classroom at a particular time, these categories and operations go well beyond a particular situation. We ought to turn, therefore, to the longer term problems of the organisation of memory.

THE STRUCTURING OF LONG-TERM EXPERIENCE

Most of us are familiar with the situation of failing to remember something which on some later occasion we can recall quite easily—or which we can recognise as soon as somebody suggests it to us. In some sense the information is stored in memory, but it is not accessible. This everyday experience might mean merely that some memories were weaker than others and that, although detectable by sensitive means, they require some extra push to bring them above the threshold of recall. Laboratory experiments, however, confirm the everyday experience, and show that it is not to be explained away in such a simple fashion. To take just one example, an experiment by Broadbent and Broadbent (1975) measured performance for recognition of words drawn from a previously presented list. This was compared with another condition in which the experimental subjects first tried to recall the words from the list. As might be expected they forgot some words. Performance at recognising these 'forgotten' words was, however, just as good as performance in the other condition. The forgetting of a word, therefore, does not mean that it is weakly stored within the nervous system : it must be that the stored trace is difficult to find under the conditions of the test of recall.

In terms of an analogy, suppose you lend a book to a friend : there may then be two quite different reasons why he does not return it. The first reason is that you may not have written your name in it, so that when he happens to come across the book on his shelves there is nothing to tell him who is the owner. The other possible difficulty is, however, that he has a lot of books in his house, that they are arranged in random order, and that consequently he cannot find your book although he knows that it is yours. The first of these is analogous to a test of recognition, in which the item is presented but the problem is to decide whether it has occurred previously. The second situation is analogous to the test of recall, in which the problem is to find the item which occurred on (belongs to) a particular occasion. One advantage of this analogy is that it makes clear that the storage of information is much less likely to cause trouble if certain steps are taken at the time of storage : if your friend thinks he is likely to meet the book and wonder whose it is, he should write your name in it. If on the other hand he thinks he is more likely to meet you and to be asked for the book, then he should keep his books in an organised fashion, such as alphabetically by authors or with borrowed books in a special place or with books on a certain topic in a certain place, or something of that sort. If we return to the case of human learning, the ability to make later use of one's experience depends on the encoding or organisation of material at the time it arrives, and this has been the topic of many recent experiments.

One common theme, for example, has been the advantage for memory of subdividing experience into a number of related clusters or groups. Thus, if one presents lists of words in which several words belong to one category such as being parts of the body, while others belong to other categories such as being names of animals, types of transport, etc., it has been known for some time that the words in the same category tend to be recalled together rather than separately (Bousfield, 1953). Tulving and Pearlstone (1966) showed that recall in such a situation was helped by providing the experimental subject with the names of the categories to which the words belonged. It is important that the categories in which the material is divided are genuinely imposed by the learner rather than merely being present in the material. Mandler and Pearlstone (1966) showed that words arranged according to the subject's own organisation were more readily learnt than words in a scheme devised by the experimenter. Nevertheless, the way the material is presented may make it much easier or much harder for the learner to structure it in useful ways : one of the most dramatic examples is a study by Bower *et al.* (1969). In this case, lists of words were given which could objectively be divided into a number of sub-categories, and these in turn could be grouped together into larger categories, making a tree or hierarchy. For example, living things could be divided into plants and animals, the latter into vertebrates and invertebrates, and so on. It was much easier for learners to reproduce the material if they had had the classificatory scheme available at the time of the presentation, even though it was not given at recall.

This hierarchical classification of the material is very much the sort of system used in many libraries and can send us back to our analogy of somebody finding a book. It would seem that somebody who learns new material with a tree-like classification in mind has the same kind of advantage as the book-borrower who places the book on a shelf reserved for the modern English subsection of the bookcase devoted to poetry, which is in the room in which he keeps non-technical books. From the educational point of view, the point is once again that routine learning by drill is unlikely to give satisfactory performance later, but that equally it will not be useful to leave information about in a haphazard way for learners who have no strategy of analysis for classification of that which is being presented to them. As the experiment of Bower and his colleagues shows, such a procedure does not arise by the light of nature : it has to be implicit in the presentation, or possibly trained as habitual sets of operations to be performed on anything presented for learning.

Useful though it is, the hierarchical method of organisation is not the only possible one. Much of the mental furniture of Western man consists of tree-like structures of subsets and supersets. Collins and Quillian (1969) even suggested that such a method of organisation could be revealed in the time taken to check the truth or falsehood of certain statements. For example, it takes much less time to accept the statement that 'canaries are yellow' than 'canaries breathe,' where the former is a specific property of canaries but not of all birds, while the latter is a property not merely of birds but of other kinds of animal as well. This is the sort of effect one might find if knowledge was stored inside the brain in a branching tree, so that it took less time to check the properties of a canary than it would take to go upwards by even one step and check the properties of all birds. However, this striking conjecture does not seem to be generally true, and probably applies only to particular kinds of material (Landauer and Meyer, 1972). Hierarchical organisation is only one of the possible methods : the property of being yellow is not specific to a small subset of birds, but is shared with bananas which are not animals at all, or with gold which is not even living. There is, therefore, a different possible classification in terms of primary colours, exact shades and degrees of saturation and so on, cutting altogether across the kind of classification we mentioned earlier. A true representation of everything that is said or remembered about the world would be a network rather than a tree, so that all the yellow objects would be connected to a common point in the network, as well as all the birds connected to a different point. Furthermore, if the model of the world inside our heads is accurate, it would need to contain the nature of association in each case : it is not enough to know that ' man ' and ' dog ' are both associated with ' bites,' because it is also very important to know whether the man is the agent or the patient of the verb, the biter or the bit. As a result, a complete network of properties and relationships will become very complex indeed. Examples are given by Lindsay and Norman (1972), and it can be seen that a network representing the information contained in a single sentence may spread over a whole page. In practice, the classifications used by human memory have to be rather simpler than this : they may be hierarchical, or cross-classified by, for example, colour and species in the example given above ; or they may be by a single dominant link from each item to one other without the nature of the link being the same in each case (canary-sing-opera-conductor-bus). Each of these takes one set of paths out of the full possible network which represents the world, and uses it as a way of linking parts together—just as our book-borrowing friend is free to arrange his books in a number of different possible ways. As we have already said, he could have a hierarchical arrangement in which modern English poets were a subset of poets in general, and poets a subset of non-fiction as a whole ; or he could have a cross-classified system in which poems were on one shelf and novels on another but modern English books were on the left-hand end of the shelf in each case and eighteenth century French at the right-hand end of the shelf. He could even, as many of our friends do, put each book next to another which seems to him relevant, although the whole shelf may then have a very bizarre look.

As yet, we know little of the relevant advantages and disadvantages of organising memory—nor indeed of organising one's own books ! Almost certainly, there is no single best method and the particular circumstances of a person will determine the approach they would be best advised to use. From the point of view of an educationalist, however, the main fact which should be borne in mind already is that fresh knowledge will always be entered into an organisational system which already exists, even though it may be primitive. If one presents material for learning in a fairly free situation and notices the manner in which people set about learning it, extreme cases can be found in which an individual works always on the basis of a single property at a time, so that he analyses what he is learning in a piece-by-piece way similar to that involved in a hierarchical system. Others on the other hand work with combinations of properties, as if they were using a criss-crossing method of organisation. Most important of all, each kind of individual learns better with material designed on his own principles than he does on material arranged in accordance with the opposite principle (Pask, 1972 ; see also Chapter 6 in Broadbent, 1973). Different teaching methods are appropriate for different people : it is particularly important to realise this, because each of us find it intuitively unreasonable that a method of presentation he himself finds hard to understand may be positively easier for somebody else. At the level of anecdote, I can well recall a child taught history in terms of striking human incidents and personal drama, on the grounds presumably of adding interest, who complained bitterly that nobody had ever told her whether Nelson put his telescope to the blind eye before or after King Charles hid in the oak tree.

To summarise thus far, information of value must receive attention if it is to be stored : and it must also be categorised or processed so as to place it in an organised structure of knowledge. Because that structure already exists at least in part, different teaching methods will be appropriate for different individuals. All that we have said thus far, however, seems to refer to the knowledge which the child acquires, rather than to the development of the powers of the individual. It could well be argued that it is the latter rather than the former which is the chief business of education, and that development consists of fresh expansion of abilities rather than dead storage of facts. This would, however, be a false antithesis, and one of the best ways of seeing this is to consider some of the recent theoretical approaches to a popular problem in development, the attainment of the ability to perform tasks of the kind designed by Piaget.

PRODUCTION SYSTEMS AS A THEORETICAL APPROACH TO BEHAVIOUR

The next part of this paper will be based upon theoretical ideas put forward by Newell and Simon (1972). In their original presentation they were applied to adult problem-solving, but they have also been used as an approach to understanding the behaviour of children, in particular by Klahr and Wallace (1972). Because the ideas are rather unfamiliar, it is probably easiest to start by putting them forward dogmatically and without qualifications, so as to follow our own earlier recommendations of securing attention to essential rather than irrelevant information ! Some possible lines of criticism can be put later, but the strength of the approach to be described is that it dissects mental functions into separate components, each of which is extremely simple. This advantage would not be affected by any of the criticisms which will be made later.

FIGURE 3

A Set of Rules for Changing Gear.

LTM

Engine screaming ⟶ CHANGE UP

Engine labouring ⟶ CHANGE DOWN

CHANGE UP + lever moved + pedal up ⟶ delete CHANGE UP

CHANGE UP + lever moved + pedal down ⟶ pedal up

CHANGE UP + pedal down ⟶ move lever

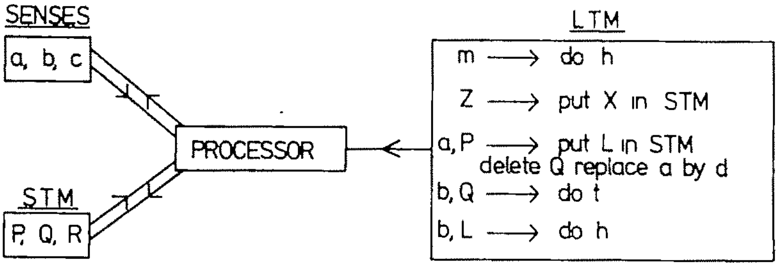
CHANGE UP + pedal up ⟶ pedal down and ease engine

CHANGE DOWN + - - - - - - - - - - -

Lower-case words represent events in the outside world, and capital letters those in short-term memory. When the events on the left-hand side of a rule are observed, the action on the right-hand side is taken. Thus, if the engine is screaming, the driver puts 'CHANGE UP' into short-term memory. If two rules are both applicable, the one higher in the list controls action.

FIGURE 4

A TYPICAL ' PRODUCTION SYSTEM ' THEORY OF COGNITIVE PROCESSES.



Lower-case letters represent events in the outside world, and capital letters those in short-term memory (STM). In the long-term memory (LTM) there are rules indicating the action to be taken when certain combinations of outside and STM events occur.

Let us start by thinking of a man driving a car, and trying to devise a set of rules which could control his gear-changing, and which might be supposed to rest in his long-term memory (see Figure 3). Each of the rules in long-term memory consists of two parts : on the left-hand side of the rule there are specified conditions under which that rule is to be used, while on the right-hand side are the actions of the processor which are to be performed if the rule is in fact put into operation. The conditions are events either in the outside world or in short-term memory : in our diagrams the former are represented by lower case letters and the latter by upper case letters. When uncertain what to do, the man examines the list of rules successively one after another, until the first time a rule is met whose conditions correspond to the actual state of affairs. Thus, with the set of rules in Figure 3, the man would do nothing until his senses informed him that the engine is either labouring or screaming. If it is screaming, the symbol 'CHANGE UP' will be inserted in the short-term memory When the list is re-examined, normally speaking the first rule which will then apply will be the sixth, which will result in pressing the clutch pedal and easing the accelerator. This in turn will make the fifth rule applicable, which will produce a movement of the gear lever : this will make the fourth rule applicable and result in the release of the clutch pedal. Only at this stage will the symbol ' CHANGE UP ' be removed from short-term memory by the operation of the third rule, and all will continue as before. We have, however, taken only one possible sequence of operations : if for some reason there was a difficulty half-way through the sequence, later stages would not occur until the hitch had been removed, and correspondingly early stages could be skipped if the required actions had already taken place. At the end of the sequence, if changing up resulted in the engine labouring, changing down would occur automatically and so on. Hence, a very complicated network of operations can be represented by a rather short set of rules.

Now let us think more abstractly, not of a man in a car, but of a mechanical system which would act in this same flexible and adaptable way. Essentially, we are supposing a system of the kind shown in Figure 4. This diagram bears close resemblance to Figure 1. Once again, it is a block diagram containing four boxes, the senses, the short-term memory, the long-term memory and a ' processor.' The latter acts in a fixed and mechanical fashion, to change either the symbols in short-term memory or the state of the outside world and, therefore, of the senses. Exactly what it does is decided by the content of long-term memory, which consists of the series of rules.

In the state of affairs in Figure 4 neither of the first two rules would go into action, but the third would. It would result in the deletion of an item from short-term memory and its replacement by another, and also in a change in the outside world. After this, the list of rules would be examined again. and thanks to the changes which have been made the third rule would no longer operate, and only the fifth would apply.

A system of this kind fits very well with the kind of thinking suggested by recent experiments on human behaviour. Indeed, its resemblance to Figure 1 makes this scarcely surprising. When running, however, it will go through a series of operations which might be described by a flow chart of the type shown in Figure 2. It needs no ' little man ' to control it, and it has the merit that it will represent a very large and elaborate flow chart within a relatively small system of rules.

As the next step, let us consider the kind of rules which would be necessary for the system to undertake simple tasks which children can perform, such as counting the number of objects of a certain shape. We will suppose that there are only four positions in which shapes can appear : for convenience, let us suppose that there are squares in positions 1 and 3, and a circle in position 2. Part of a suitable set of rules for long-term memory appears in Figure 5. This particular set of rules will allow the system either to count squares, to count circles, to copy the symbols or to produce a reversed copy of the symbols, depending upon the particular instructions which are residing in short-term memory. Let us trace through, as an example, the sequence of events if the system is instructed ' COUNT SQUARES ': the contents of short-term memory, and the action taken, at each successive stage in time are shown in Figure 6.

FIGURE 5

A SET OF LTM RULES FOR COUNTING, COPYING, OR COPYING IN REVERSE FOR A VISUAL DISPLAY CONTAINING THREE POSITIONS, EACH HOLDING EITHER A SQUARE OR A CIRCLE.

LTM

POSITION 4 BEING EXAMINED ——→ STOP

STOP+COUNT ——→ Report total

CONTINUE ——→ Add one to ' POSITION BEING EXAMINED,' delete CONTINUE and detect

COUNT+no memory of total ——→ Remember TOTAL=0

$\left.\begin{array}{l}\text{COPY}\\ \text{REVERSE}\\ \text{COUNT}\end{array}\right\}$ +no memory of POSITION BEING EXAMINED ——→ Remember POSITION ONE BEING EXAMINED and detect

Square detected+COPY ——→ draw square and CONTINUE

Circle detected+COPY ——→ draw circle +CONTINUE

Square detected + REVERSE ——→ draw circle+CONTINUE

Circle detected+REVERSE ——→ draw square+CONTINUE

Square detected+COUNT SQUARES ——→ Add one to TOTAL and CONTINUE

Circle detected+COUNT SQUARES ——→ CONTINUE

Square detected+COUNT CIRCLES ——→ CONTINUE

Circle detected+COUNT CIRCLES ——→ Add one to TOTAL and CONTINUE

The first step is that, by the operation of rule four, short-term memory records ‘TOTAL=ZERO.’ The next event that occurs is that by the operation of rule five, the system examines position 1 to see what symbol is present. By the operation of rule ten, this results in ‘TOTAL BEING INCREASED BY ONE,’ and the next position being examined. After all three positions have been examined, the system will stop with a total equal to two, which is, of course, the number of squares present. The same system of rules would have resulted in a quite different sequence of operations and a different result if the symbols presented to the senses had been different, or if the instructions had been different.

FIGURE 6

The Events Which Occur if the System of Figure 5 is Told to ‘Count Squares’ and Shown a Square, a Circle and a Square.

SUCCESSIVE TIMES	STM			ACTION
	INSTRUCTION	POSITION BEING EXAMINED	TOTAL	
1	COUNT SQUARES	—	—	store ‘TOTAL=0’
2	COUNT SQUARES	—	0	store ‘POSITION 1 BEING EXAMINED’ and detect (finding a square)
3	COUNT SQUARES +CONTINUE	1	1	Add one to POSITION delete CONTINUE and detect (finding a circle)
4	COUNT SQUARES	2	1	Restore CONTINUE
5	COUNT SQUARES +CONTINUE	2	1	Add one to POSITION delete CONTINUE and detect (finding a square)
6	COUNT SQUARES	3	2	Restore CONTINUE
7	COUNT SQUARES +CONTINUE	3	2	Add one to POSITION delete CONTINUE and detect (finding nothing)
8	COUNT SQUARES	4	2	STOP

The ability to count objects of a certain type, which we have just been discussing, is obviously one of the elements in successful performance of a task such as the ‘class inclusion’ task of Piaget. That is, presentation of a set of objects all of which possess property A, while some of them possess property B and others do not: with the question “Are there more A or B?” Wallace (1974) has generated sets of rules which will perform this task, and which will fail if certain rules are missing. He has also compared the behaviour of actual children before and after the stage of successful performance of the task, in order to observe similarities and differences between their behaviour and that of the theoretical model. Certain elementary forms of the theoretical model can easily be disproved : for example, if the child were to look at each member of the subset in turn, add one to a ' Total ' and mark the item as counted it could succeed in the task, and it would fail if something went wrong with the marking of items as having been counted already. However, a system of this kind would either perform quite well even before it reached the stage of perfect performance (because if it happens to count the larger set first, it would get the right answer) or else it would show a bi-modality of response times. Neither of these findings is in fact observed, and so this over-simple theory can be disproved.

When, however, a more complex form of model was constructed, it justified itself by displaying behaviour which was unexpected and which nevertheless did appear in the performance of actual children. In the model finally adopted by Wallace, the system would search for an example of the class which was to be counted, and then count all similar objects. There is a deliberate ambiguity in the word ' similar.' Clearly, the system could either count objects similar to the example in every way, or objects similar only in the particular feature which led to the example being chosen in the first place. Suppose we imagine a child trying to count the squares when faced by an array of eight red and four black squares Whichever object is picked first will be similar to all the other eleven in respect of its squareness, but to a smaller number in respect of its colour. Thus, until a rule for picking objects can be based on only one sensory dimension, the child would fail the task. This implies failure at the apparently simple task of merely counting the squares, without the extra complications implicit in the usual question "Are there more squares or red objects ? " Most attempts to explain this process of development assume that counting as such is well established. Yet, in fact, in Wallace's experiments any child who failed the usual class inclusion question also failed the apparently simple task of counting the larger class of objects. Training could be given in this particular function, by presenting ' objects ' which were actually patches of colour seen through separate shaped masks, so that one dimension could be separated from the other very easily as an illustration of identity. After such instruction, some children could succeed in the original class inclusion task with the extra elements of repeated counting and comparison.

This form of theorising has therefore been justified, at the least, by its suggestions of a novel observation which sheds rather different light on the process of development observed by Piaget. We will return shortly to the general implications of this system for education ; before doing so, it may be appropriate now to point out some of the limitations.

Criticisms of the ' Production Systems ' Approach.

It is unlikely that anybody would wish to challenge some features of this approach, although they may not have been stated so clearly before the work of Newell and Simon. It does seem highly probable that any adequate model of a child's behaviour will include some long lasting rules of behaviour common from one situation to another. It will also include some temporary or short-term representation of the fact that the child is, at this instant, counting rather than copying : this kind of memory must last throughout the duration of a task, but disappear when the task is complete. Similar arguments apply to the short-term memory of the total so far counted, or the objects which have been counted previously. Again, there would be widespread agreement about the existence of some ' processor ' which creates the short-term memory of the total counted

(' says it to oneself') and undertakes similar functions. There are, however, a large number of details in the model which are inessential, and could quite well have been different. To take the least important first, the processor is supposed to undertake only one action at a time—as do the processors in contemporary computers. This is almost certainly too simple an account of what happens in human beings, since our knowledge of the structure of the brain certainly suggests that some low level processing goes on in parallel with that occurring at higher levels. The *extent* to which parallel processing can occur throughout the system, is, however, somewhat controversial. Even the evidence suggesting a limited capacity for processing information does not mean that processing must be entirely serial, if the quantity of information involved is relatively small. However, for rough practical purposes one can accept a serial processor as an approximate model of what happens in human beings (Broadbent, 1971).

A little more serious is the fact that this model uses only one kind of short-term memory : Wallace, in fact, already modifies this by supposing separate short-term memories for different kinds of information and from the data of cognitive psychology one would certainly expect partially separate memory storage for visual and for auditory information, for semantically encoded information, and so on. One could perhaps put this point as sharply as possible by saying that, in this model, it should be possible to get a spoken report of the goal which is controlling behaviour at any moment ; but this seems a little doubtful in fact.

Perhaps the most arbitrary feature of this approach lies in the procedure adopted for deciding that one rule rather than another shall control behaviour. As described, the system searches one after another through the rules in long-term memory, and stops at the first one which meets the case. When examining each rule, it looks simultaneously at everything present to the senses and in short-term memory. Yet one could imagine a system working on exactly the opposite basis, which considers simultaneously all the rules in long-term memory, but looks one after another at each item in short-term memory and at each object in the outside world, stopping the first time it meets a condition which satisfies the rule. Even more frightening, a little introspection suggests that human beings may work on one basis in some situations and on another basis in other situations : what would be the rule for deciding which procedure to adopt ?

Educational Implications of the Approach.

The foregoing criticisms have been given because they are bound to strike any careful reader as being arbitrary features of the theory. It is important, therefore, to realise that they do not affect the main strength of the approach, in splitting the flow of behaviour into operations, temporary states of the system and continuing rules of action. In particular, this kind of analysis shows that a general mental capacity can depend logically on a particular operation, whose absence may be as fatal as the absence of a keystone in an arch. In the particular case of the class inclusion task, the principle emphasised by Wallace draws a link with earlier sections of this paper, because it is a principle which selects the next item to be processed—shall it be an item similar to the last item in one feature or in many, and shall similarity to the last item be in the same feature as the similarity to the last but one ? The principles of organisation and classification, which we saw at play in memory, appear also in intellectual operations, and their absence can make such operations impossible. Once again, therefore, the lesson of cognitive psychology is opposed both to teaching by routine drill on the one hand and also to a passive exposure of the child to material on the other.

GENERAL CONCLUSIONS

To summarise again, the message of cognitive psychology is :

(*a*) Only some of the information presented will receive attention, and if this is not decided deliberately it will certainly be decided by chance factors.

(*b*) When something is noticed in the environment, it can be processed or encoded in a number of ways, and the particular processing which takes place will decide whether the effects are long lasting or transitory.

(*c*) If the effects are to be enduring, then the information must be organised at the time it is stored, in such a way that a clear path leads to it from the likely situations in which it may be needed.

(*d*) Possible methods of organisation range from the clear (but rigid) system of hierarchical subdivision to the flexible (but confusing) network of private association.

(*e*) By the time a teacher meets any child, some system of organisation will probably already exist, and it is likely that fresh information will be best assimilated through the system which is already tentatively established.

(*f*) The achievement of formal intellectual skills probably depends on the availability of a set of general procedures or rules, each of which comes into play as the momentary external and internal circumstances make it appropriate. Because of the complex interplay of the various factors involved, the absence of one link in the chain will not be obvious, but will make certain formal processes of thought extremely difficult, and it is, therefore, necessary to look for such links. Particularly crucial ones will involve the organisational and selective factors already mentioned.

The study of cognition was, of course, a classic interest of psychology, and in some ways current concerns are taking up again older problems with new sophistication. It is perhaps appropriate, therefore, to remember the classic doctrines of ' faculty psychology ' long derided in modern texts. It was indeed an error to suppose that the study of Latin and Greek would develop intellectual ability regardless of the way in which it was taught, or of the specific field to which the intellect was to be applied in later life. This does not mean, however, that the school environment is merely a cafeteria in which an autonomously developing intellect picks whatever suits its needs, nor that learning consists purely of elementary associations of stimuli and responses. Rather, the lesson of cognitive psychology is that each of us acquires during life certain strategies of encoding the outside world, of organising memory and of proceeding from one step in an operation to the next, and that these may be highly general in their later use. The successful teacher, of course, has always known this, but in standing out for the middle ground between mechanical drill on the one hand, and the abandonment of all positive teaching on the other, he can now claim the support of contemporary cognitive psychology.

REFERENCES

BOUSFIELD, W. A. (1953). The occurrence of clustering in recall of randomly arranged associates. *J. gen. Psychol.*, 49, 269-273.

BOWER, G. H., CLARK, M. C., LESGOLD, A. M., and WINZENZ, D. (1969). Hierarchical retrieval schemes in recall of categorised word lists. *J. verb. Learn. verb. Behav.*, 8, 323-343.

BROADBENT, D. E. (1958). *Perception and Communication*. London : Pergamon.

BROADBENT, D. E. (1971). *Decision and Stress*. London : Academic.

BROADBENT, D. E. (1973). *In Defence of Empirical Psychology*. London : Methuen.

BROADBENT, D. E., and BROADBENT, M. H. P. (1975). The recognition of words which cannot be recalled. In DORNIC, S., and RABBITT, P. (Eds.), *Attention and Performance V*. London: Academic.

COLLINS, A. M., and QUILLIAN, M. R. (1969). Retrieval time from semantic memory. *J. verb. Learn. verb. Behav.*, 8, 240-247.

IMAI, S., and GARNER, W. R. (1965). Discriminability and preference for attributes in free and constrained classification. *J. exp. Psychol.*, 69, 596-608.

KLAHR, D., and WALLACE, J. G. (1972). Class Inclusion Processes. In FARNHAM-DIGGORY, S. (Ed.), *Information Processing in Children*. London : Academic.

LANDAUER, T . K., and MEYER, D. (1972). Category size and semantic memory retrieval. *J. verb. Learn. verb. Behav.*, 11, 539-549.

LIGHT, L. L., and CARTER-SOBELL, L. (1970). The effect of changed semantic context in recognition memory. *J. verb. Learn. verb., Behav.*, 9, 1-11.

LINDSAY, P. H., and NORMAN, D. A. (1972). *Human Information Processing*. New York : Academic.

MACCOBY, E. E., and KONRAD, K. W. (1967). The effect of preparatory set on selective listening : developmental trends. *Monog. Soc. Res. Child Develop.*, Vol. 32, No. 112.

MANDLER, G., and PEARLSTONE, Z. (1966). Free and constrained concept learning and subsequent recall. *J. verb. Learn. verb. Behav.*, 5, 126-131.

MORAY, N. (1959). Attention in dichotic listening : affective cues and the influence of instructions. Q. *J. exp. Psychol.*, 11, 56-60.

MORTON, J. (1967). A singular lack of incidental learning. *Nature*, 215, 203-204.

MOWBRAY, G. H. (1953). Simultaneous vision and audition : the comprehension of prose passages with varying levels of difficulty. *J. exp. Psychol.*, 46, 365-372.

NEISSER, U. (1967). *Cognitive Psychology*. New York : Appleton-Century-Crofts.

NEWELL, A., and SIMON, H. A. (1972). *Human Problem Solving*. New York : Prentice-Hall.

PASK, G. (1972). Learning strategies and individual competence. *Int. J. Man-Machine Studies*, 4, 217-253.

TULVING, E., and PEARLSTONE, Z. (1966). Availability *versus* accessibility of information in memory for words. *J. verb. Learn. verb. Behav.*, 5, 381-391.

TURVEY, M. T. (1967). Repetition and the preperceptual information store. *J. exp. Psychol.*, 74, 289-293.

WALLACE, J. G. (1974). A study of concrete operational thinking employing computer simulation techniques. *Report to SSRC on Grant HR* 1090/2.

WINOGRAD, T. (1972). *Understanding Natural Language*. Edinburgh : University Press.

(*Manuscript received 7th February*, 1975)

Br. J. educ. Psychol., 45, 177-188, 1975

PREDILECTION FOR GENERALISING AND ABSTRACTING

By E. A. PEEL

(*School of Education, University of Birmingham*)

SUMMARY. This study reports the development of a sentence preference test designed to contrast tendencies to abstract, generalise and particularise in thinking. Certain terms carry both general and abstract meanings. Using such words in sentences bringing out these differing meanings provides an indication of readiness to use abstractions and generalisations. A variety of samples of secondary school pupils in the age range 11+ to 16 years were used to provide evidence of reliability and validity, and also to trace changes with age of responses to the sentences. The results confirm earlier work with a word test which showed that tendencies both to generalise and abstract increase during adolescence.

The potentialities of the technique are discussed in relation to language usage and comprehension and to higher level thinking processes required during secondary school and afterwards in higher education.

GENERALISING AND ABSTRACTING

As the pupil and student progress towards higher levels of learning, we assume a growing tendency to utilise abstract verbal notions. The abstract term is a powerful symbol enabling the thinker to systematise knowledge in a precise way and to express hypotheses succinctly. At the same time, however, like all powerful tools, abstraction can also be misused, as when people bandy about words like democracy and equality without considering what the abstractions embody. But for all that, without general and abstract terms, verbal discourse would be very tedious and limited indeed.

When we consider studies of classification and concept formation, we often find the notions of generalising and abstracting used in an equivalent sense, for example, Goldstein and Scheerer (1941) used the phrase ' abstract attitude ' to indicate how far a patient could generalise the class of reds to include pinks. But there is a possible distinction between two aspects of forming a concept, by which extending the array of exemplars might be called generalising, and formulating the rule for including instances would require an act of abstracting. Such a distinction underlies the purpose of the present investigation. Long ago, Welch and Davies (1935) drew a distinction between undissociated and dissociated abstractions as in the instance

The cherry is red (undissociated)
redness (dissociated).

This differentiation is significant to the present study in being near to the distinction which will be made respectively between generalising and abstracting. Furthermore, they pointed out that the undissociated is genetically prior to the dissociated abstraction. In the terms of the present paper, psychologically, generalising is prior to abstracting.

A most explicit distinction between generalisation and abstraction stems from the penetrative ontological discussion by Quine (1953) on the differences between general and abstract particular terms, as between, say, *square* and *squareness*. These notions were later extended by Braine (1962) to a critique of certain explanations of young children's reasoning. Following Braine's example (*op cit.*, page 50) we might use the term generalisation for such transitive subset inclusions as

Fido—dog—quadruped—animal (increasing degrees of generalisation)
and abstraction for such intransitive class membership relationships as

Fido—dog—species—identifying principle (heightening levels of abstraction).
Finally, Lunzer (1963) pulled together these notions in his informative discussion of the connection between concept formation and abstraction.

In what follows the basic distinction between a concrete-general and an abstract-particular is embodied in language material used for testing purposes.

Often the abstract-particular is distinguished from its related concrete-general term by a suffix marker such as *ism*, as for example, in

Despotism curtails individual freedom
Henry VIII was a *despot*

but there is also a widespread use of the same word for each usage, as in

' Upon Westminster Bridge ' is a sonnet
The sonnet is a poetical form.

Here, *sonnet* is a homonym in a very intimate sense, symbolising both a concrete-general and an abstract-particular. Similarly the ambiguity of the sentence *I hate fighting* may be attributed to the two meanings of the word *fighting*, that is, the concrete-general (the act of fighting) and the abstract-particular (the notion of fighting). Such ambiguity may provide us with a tool for assessing a person's tendency to generalise or abstract.

The problem is to discover to what extent people tend or prefer to use either of these two meanings, the concrete-general or the abstract-particular (for short, the general or the abstract). The technique first devised (Peel, 1971a, 1971b) to measure these tendencies or preferences utilised the homonymous nature of certain general terms as in the following test item. The subject was asked to underline the word or phrase among the five alternatives given on the right, which for him most appropriately completed the row of three terms.

Jamaica	island	1. climate
		2. geographical notion
		3. Bermuda
		4. bathing
		5. feature of the earth's surface.

The testee commences with the concrete-particular, *Jamaica*, is led to the concrete-general terms, *island*, at which stage its alternative meaning becomes possible, that is, the abstract-particular (the notion of island). Such a shift of meaning from concrete-general to abstract-particular is what in the present context is meant by abstracting and it is considered to have taken place if the subject selects choice 2, *geographical notion*. If he selects choice 5, he is held to be thinking of *island* in the concrete-general sense. These alternate meanings of the term *island* are brought out in the sentences

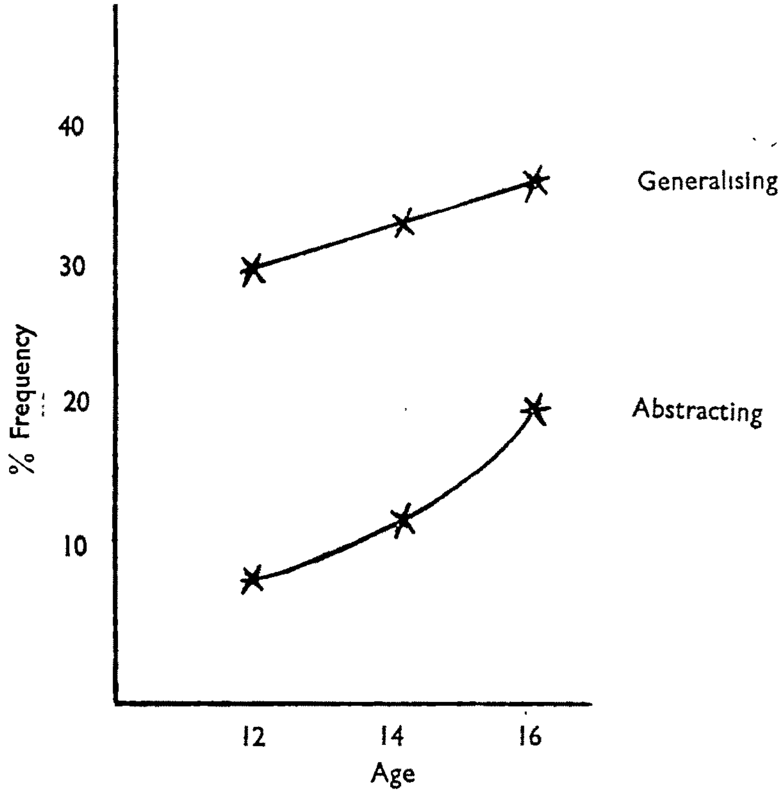
concrete-general : An island is a feature of the earth's surface.
abstract-particular : The island is a geographical notion.

If the subject selects any of choices 1, 3 and 4, it is assumed that he is making a meaningful, but not logical connection, between *Jamaica*, *island* and one of the three words, *climate*, *Bermuda* or *bathing*.

A test of 30 such items was compiled. The results from it, based on testing 63 12-year-old, 66 14-year-old and 70 16-year-old secondary school pupils (Peel, 1971a) demonstrated a rise in tendency to abstract with age, a less steep rise in tendency to generalise and a fall in other responses. The tendencies with age to abstract and generalise are shown in Figure 1.

FIGURE 1

AGE TRENDS IN THE TENDENCIES TO ABSTRACT AND GENERALISE.
(from Peel, 1971a)



When considering this first test, it was noted that its form, in leading from a concrete-particular to a concrete-general, might predispose a person towards generalising, rather than abstracting. Subsequently, three training items were utilised, giving practice respectively in generalising, abstracting and alogical association responses. These practice items, following the particular rubric chosen in which attention is drawn to relations between terms, might overstructure the test situation so that it leads to a measure of logical *ability* rather than an instrument assessing *predilection*. Also the selecting of the three other responses in the multiple choice was rather haphazard and this wasted a possible source of further information about the tendency of the subject to particularise as revealed in choice 3 of the examples, namely, *Bermuda*. Selection of this response would suggest a predilection for choosing the *members* of the island class, including Jamaica, Bermuda, etc. Lastly, it was considered desirable to try the test in a sentence preference form which might provide a more natural lead to the cognitive preference available.

A sentence preference test.

The sentence preference test consisted of 21 items, each comprising a set of four sentences. Each sentence contained the same key test word, a term being associated in all with four meanings involving the notions of abstraction (A), generalisation (G), particularisation or membership (M) and partial meaningful association with other terms in the sentence (P), as in the following examples. (The full test with key is shown as Appendix 1).

The *motorway* commends itself to many as an idea for improving motor travel. (A)
The *motorway* is a civil engineering construction. (G)
The *motorway* system includes M1 and M2. (M)
There is a *motorway* junction near Bristol where M4 and M5 connect up. (P)

Ireland is an *island.* (M)
An *island* is a feature of the earth's surface. (G)
An *island* is often rocky. (P)
The *island* is a geographical notion. (A)

It is important that the four sentences in each set are made as normal as possible (in terms of language usage). Sometimes it is in sentences involving the abstract usage of the key word in which this is most difficult to do, as in the following items :

The idea of a *helper* is a human ideal as old as written records. (A)
The good Samaritan was a *helper* of an injured traveller. (M)
A *helper* is an honest man and a kindly individual. (G)
A *helper* is needed for caring for the needs of old people. (P)

and :

Mars is a *planet.* (M)
A *planet* is greatly interesting to an astronomer. (P)
The *planet* is an astronomical class. (A)
A *planet* is a heavenly body. (G)

The subject is asked to select the sentence, out of each four, which has the greatest significance for him and the rubric is carefully chosen to elicit preference or predilection (see Appendix 1).

The 21 key concept terms, around which the sets of four sentences are built, cover a wide range of meanings on the assumption that cognitive preferences will be tested rather than knowledge in particular fields. The concepts are : city, poem, playing, motorway, justice, helper, adolescent, creature without a backbone, island, object moving under the pull of gravity, Polar region, plane surface, aggression, team of eleven players, planet, shop, painter, one day and the next day, savings account, plumber, Frenchman. The response to each item must be either A, G, M or P, and the total responses in these four categories across the whole test provide scores on these different response tendencies.

The test has been used in studies of cognitive development across the years of adolescence and also in studies of cognitive style. In these researches interest has centred mainly on the use of the test as a measure of tendencies to abstract (A) and generalise (G) and, to a lesser extent, to particularise (M).

Application of the test.

Before the evidence for systematic trends by age in abstracting and generalising was considered, the reliability of the test was calculated. The results of its item and analysis were examined and correlations set out revealing the validity of the test as measured against certain cognate mental capacities.

(*a*) *Reliability.*

Evidence of reliability was obtained by using split-half correlations derived from two independent populations, with the following results : A scores, r=·736, N=126, SE=·041
G scores, r=·864, N=200, SE=·018.

(*b*) *Item analysis.*

The form of the test, involving four alternative possible responses, A, G, M and P, poses some interesting problems and suggests some novel possibilities in methodology. Let us start by setting out the results of giving the test to two independent populations, L and B, consisting respectively of 218 and 200 secondary school pupils ranging from the first to the fourth years. We count the total number of A, G, M and P responses to each of the 21 test items in each population and then convert the frequencies to percentages of 218 and 200, respectively, to make comparison possible. These percentage frequencies are set out in Appendix 2, alongside the relevant concept term involved. We then assume that an item would produce consistent responses if there was some similarity in the pattern of percentage responses in the two sets of A, G, M and P categories. This requirement is well brought out by considering, in Table 1, the percentage frequencies for items 2 and 3, involving the notions, respectively, of *poem* and *playing*. The data given show a high degree of consistency.

TABLE 1

PERCENTAGE FREQUENCY OF DIFFERENT RESPONSES TO TWO CONCEPTS.

Item*	Concept	Percentage frequency of responses							
		From population L				From population B			
		A	G	M	P	A	G	M	P
2	Poem	17	54	9	21	15	51	11	24
3	Playing	14	6	15	65	19	7	18	57

* See Appendix 1 for full details of items.

As a first crude display of the consistency of the test as a whole, we may set out those items which produced in both populations a common highest and a common lowest percentage frequency and identify the relevant response categories. Thus, in the above two items, G is the common highest frequency in 2, P is the common highest frequency in 3, whilst M and G are, respectively, the common lowest frequencies in responses to items 2 and 3.

TABLE 2

MOST AND LEAST FREQUENT RESPONSES TO ITEMS.

Item :	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Highest common frequency	.	G	P	A	.	G	G	A	G	G	.	.	G	M	M	G	G	G	P	.	G
Lowest common frequency	A	M	G	P	.	A	.	.	P	P	P	G	.	P	A	P	M	.	M	P	P

Table 2 summarises the most frequent and least frequent responses to all items in Appendix 2. In 16 items the most popular preferences coincide, in 16 items the least popular coincide and in 12 items both most and least popular coincide. In only one item is there no coincidence in responses, that is to item 5, involving the concept of *justice*. This item appears too difficult.

Justice was done when the train robbers were convicted. (M)
Justice is a basic principle underlying the continued existence of society. (A)
Justice is the punishment of a criminal. (P)
Justice is a human practice making it possible for people to live together more securely. (G)

The nature of the most and least popular responses can be obtained by counting the relevant frequencies of A, G, M and P.

We may improve on this rather thumbnail qualitative statistic by a simple device, that is to rank the percentage responses to the 21 items in each of the 8 columns in Appendix 2, by the usual method. Thus, for the Abstraction column in population L, item 4 would have rank 1, item 17 rank 2, items 9 and 11 have each rank 4½, and so on to item 5 with rank 21. We can then calculate the rank correlation coefficient between the A responses in populations L and B and do the same for G and M. When this is done, the results are as follows :

$$r_A = \cdot 802, \quad r_G = \cdot 864, \quad r_M = \cdot 840.$$

The test overall then produces consistent results whichever responses we consider, i.e., A, G or M.

(*c*) *Validity of the test.*

Assessments of the validity of the preference measures were obtained by correlating secondary pupils' A and G preference scores with their scores in tests of related cognitive functioning. These latter included verbal reasoning, judgment from textual material (Peel, 1973), use of abstract notions in the subjects' own writing (Gillie, 1957), T-unit length (Hunt, 1965),other measures of written language (Bartholomew, 1974), reading comprehension (Bate, 1966) and fluency. Results collected from several pieces of research are set out in Table 3.

TABLE 3

PRODUCT MOMENT CORRELATION COEFFICIENTS BETWEEN A AND G PREFERENCES AND OTHER INTELLECTUAL VARIABLES.

Intellectual variable	N	A preference		G preference	
		r	SE	r	SE
A.H.4	122	·392	·076	·283	·083
	88	·407	·089	·266	·099
Judgment	146	·260	·077	·086	·082
	130	·294	·080	·287	·081
Bate NFER Reading comprehension	160	·310	·071	·358	·069
Pupil's own writing :					
Subordinate clauses	100	·315	·090		
T-unit length	100	·204	·096		
Gillie abstractness	100	·212	·096		
Fluency :					
Words	100	·271	·093		
Ideas	100	·323	·090		

The results displayed in Table 3 portray significant correlation between the preference tests and tests of cognitive functioning and language usage. We may therefore have some confidence in the test as a measure of intellectual preferences.

CHANGES WITH AGE IN THE PREDILECTION FOR GENERALISING, ABSTRACTING AND PARTICULARISING

The earlier research (Peel, 1971a, 1971b) suggested that there were significant gradients with age in the frequency of A responses and in the G responses to a lesser extent. At the same time, as Figure 1 shows, the predilection for generalising was higher throughout and, as one critic suggested, this might be an artefact arising from the nature of the first form of the test which led the testee from a concrete-particular to a concrete-general before he made his choice.

Three studies, using the sentence preference test, have been made on the change of A, G, M and P frequencies with age. In the first study, 333 secondary school pupils made up of groups of individuals in the age range 11–16 years were tested. Their mean responses are shown in Table 4.

TABLE 4

MEAN SCORES AND STANDARD DEVIATIONS ON THE SENTENCE PREFERENCE TEST BY AGE (11-15 years)

Age	N	Response Categories			
		A	G	M	P
11 : 9+	93	3·72 (2·32)	6·15 (2·79)	6·40 (3·58)	4·71 (2·07)
13+	71	3·92 (2·75)	6·95 (2·56)	5·80 (3·13)	4·28 (1·88)
14+	93	4·97 (3·38)	7·61 (2·67)	4·74 (3·48)	3·65 (2·10)
15+	76	5·61 (3·80)	7·53 (2·45)	3·99 (2·94)	3·82 (2·25)

TABLE 5

STANDARD ERRORS OF THE DIFFERENCES BETWEEN THE MEANS OF DIFFERENT AGE-GROUP PAIRS (Figures in parentheses are the actual differences between the relevant means obtained from Table 4.)

Age	Standard Errors and differences between means			Significance of the differences between the means		
	11 : 9+	13+	14+	11 : 9+	13+	14+
	A responses					
13+	·406 (0·20)			*		
14+	·425 (1·25)	·479 (1·05)		2SE	>2SE	
15+	·497 (1·89)	·545 (1·69)	·559 (0·64)	>3SE	>3SE	*
	G responses					
13+	·421 (0·80)			2SE		
14+	·402 (1·46)	4·11 (0·66)		>3SE	*	
15+	·392 (1·38)	4·14 (0·58)	·395 (-0·12)	>3SE	*	*
	M responses					
13+	·525 (0·60)			*		
14+	·518 (1·66)	·518 (1·06)		>3SE	>2SE	
15+	·502 (2·41)	·502 (1·81)	·494 (0·75)	>4SE	>3SE	*

* Insignificant differences.

Consider now the significance of the differences between the means of A,G and M, across these age groups. For the differences between the means of the respective classes of responses, we have (Table 5) the values of their standard errors. Alongside we set the array of those cases in which the actual differences between the means (taken from Table 4) exceed their standard errors by at least two times.

There appears to be firm evidence of a significant trend by age upward in the case of abstracting (A) and generalising (G) and downward in the case of particularising (M).

A second investigation involved 220 secondary school boys and girls aged between 11 : 10 and 18 : 09 years. The results are set out in Table 6.

TABLE 6

Mean Scores and Standard Deviations on the Sentence Preference Test by Age (11-18 years)

Age Range	N	A	G	M	P
11 : 10—12 : 09	27	2·15 (1·96)	5·33 (2·19)	8·07 (2·81)	5·44 (2·13)
13 : 10—14 : 09	84	5·36 (3·41)	7·49 (2·77)	4·61 (3·39)	3·55 (2·15)
14 : 10—16 : 09	58	5·43 (3·75)	7·42 (2·87)	5·54 (3·91)	2·62 (1·74)
16 : 10—18 : 09	51	5·86 (3·01)	7·78 (2·52)	4·90 (3·96)	2·45 (1·65)

As will be seen from Table 6, the same trends were demonstrated as obtained in the first investigation, and we note also that the mean scores are very similar in both cases.

In a third investigation involving 160 secondary school pupils, the group was divided into two halves of 80, according to their score in the AH4 test. Pupils were equally represented in each of four age levels (20 in each AH4 group). Results of this analysis are shown in Table 7. Again, the same age trends appear and there are also clear differences associated with the low and high AH4 scores.

TABLE 7

Mean Scores and Standard Deviations on the Sentence Preference Test by Age and Ability Level.

Age	N	Low AH4 Score				N	High AH4 Score			
		A	G	M	P		A	G	M	P
12+	20	2·80 (1·69)	5·91 (2·66)	6·50 (3·63)	5·80 (1·87)	20	3·50 (1·83)	7·60 (1·66)	6·00 (3·08)	3·90 (1·81)
13+	20	3·60 (2·28)	5·70 (2·42)	6·40 (3·36)	5·50 (2·19)	20	4·90 (1·43)	8·80 (2·55)	3·60 (2·55)	3·90 (1·88)
14+	20	3·60 (1·96)	5·80 (2·42)	6·60 (3·15)	5·00 (1·79)	20	5·00 (2·75)	7·80 (1·71)	4·10 (2·33)	4·30 (1·72)
15+	20	3·90 (1·61)	7·70 (0·75)	5·00 (2·77)	4·50 (1·75)	20	5·80 (2·63)	8·20 (2·35)	4·80 (2·78)	2·40 (1·18)

There are then definite trends by age, increasing in the case of abstracting (A) and generalising (G) and decreasing in the case of particularising (M). Incidentally, we may note also that the partial association scores similarly decrease with age. Discussion of this has not been taken up since the P scores are determined once the A, G and M scores are known. The generalising (G) responses are relatively much less numerous than in the earlier research (Peel, 1971a) and it seems, therefore, that the form of the earlier rubric might have encouraged generalisation responses. On the other hand, the sentence form may have introduced difficulties concerning naturalness, which might be obviated by reverting to a simplified single term form as in the following items :

Stimulus word : city *Response from* : administrative concept
Liverpool
difficult place for parking cars
unit of local government.

Such a form is currently being tested.

IMPLICATIONS AND POSSIBILITIES

The findings imply that in addition to an increasing dependence on generalising during adolescence, there is evidence of a growing preference for abstract notions. Such findings are in line with the affirmation of at least one linguist (Hayakawa, 1949), who devised the term 'abstraction ladder' for changes in the process of higher learning and asserted that consciousness of abstracting is a sign of adulthood.

The trends in the four categories have educational and psychological significance, and, given an instrument which can reveal tendencies to use these categories as the tests appear to do, we could open up interesting research in problems of education, cognition and language development. Much of secondary and higher education demands familiarity with abstracting, both in textual comprehension and production. The highest levels of judgment and explaining are often in terms of hypotheses implying abstractions, as say, in thinking in science, geography and literary criticism. The thinking and writing required of the student at these levels calls heavily on generalising and abstracting. The test, therefore, could be used in studies of the growth of higher level learning in different school subject areas. In this connection the test could be used also as a readiness or predictive measure. It could be applied as a component in batteries of tests designed to investigate cognitive and language development, not only with school age adolescents but also with atypical groups from other cultures, possessing English only as a second language.

One might also investigate a possible link between generalising and abstracting tendencies and the readiness to explain phenomena, as opposed to merely describing them and in the equivalent urge to invoke outside ideas (Peel, 1975).

The notions underlying the testing technique can be applied to give a measure of the extent of the generalising and abstracting evident in a person's continuous writing. Such a scheme is being applied to specimens of productive writing.

Finally, an interesting psychological possibility is whether the patterns of abstracting and generalising are sufficiently stable in late adolescence and adult years to warrant their description as cognitive styles. The concept of generalising and abstracting could be compared with the concepts of equivalence of range and bandwidth.

ACKNOWLEDGMENTS—I am particularly indebted to my research students, Mrs. M. L. Bartholomew, Mrs. K. B. Lewry, Mr. F. H. Boardman and Mr. J. I. Ellis for their kindness and efficiency in applying the test with a variety of other cognitive and linguistic test material to different populations.

REFERENCES

BARTHOLOMEW, M. L. (1974). Language usage, abstraction, judgment and fluency in early adolescence. Unpublished M.Ed. thesis, University of Birmingham.

BATE, S. M. (1966). *Reading and E.H. Test* 2 ; *Comprehension.* Slough : NFER.

BRAINE, M. D. S. (1962). Piaget on Reasoning : A methodological critique and alternative proposals. IV. In 'Thought in the Young Child,' *Monogr. Soc. Res. Child Dev.*, 27, 2, 41-60.

GILLIE, P. J. (1957). A simplified formula for measuring abstraction in writing. *J. appl. Psychol.*, 41, 214-217.

GOLDSTEIN, K., and SCHEERER, M. (1941). Abstract and concrete behaviour. *Psychol. Monogr.*, 53, 1, 151.

HAYAKAWA, S. I. (1949). *Language in Thought and Action.* New York : Harcourt.

HUNT, W. K. (1965). Grammatical structure at 3 grade levels. *National Council for Teachers of English Report No.* 3. Champaign, Illinois.

LUNZER, E. A. (1963). Concept formation and abstraction. Nottingham University School of Education Unpublished Research Report.

PEEL, E. A. (1971a). Generalising and abstracting. *Nature*, 230, 5296, 600.

PEEL, E. A. (1971b). Adolescent concept formation. In *Interdisciplinary Approaches to Language.* Centre for Information in Language Teaching Reports and Papers No. 6, Paper 2, 21-29.

PEEL, E. A. (1973). *The Nature of Adolescent Judgment.* London : Crosby Lockwood Staples.

PEEL, E. A. (1975). The analysis of comprehension and judgment from textual material. *Educ. Rev.*, 27, 100-113.

QUINE, W. V. O. (1953). *From a Logical Point of View.* Cambridge, Mass. : Harvard University Press.

WELCH, L., and DAVIES, H. I. (1935). The theory of abstractions. *Psyche*, XV, 138-145.

(*Manuscript received* 21*st October*, 1974)

APPENDIX 1

SENTENCE PREFERENCE TEST.

There follow sets of four sentences, each set being concerned with the single term in *italics*. All the sentences are factually correct. Read through each set of sentences and tick the one out of the four you prefer as having the most significance for you.

Key

1. A The *city* is an administrative concept going back to Roman times.
   M Liverpool is a *city*.
   P A *city* is a difficult place for parking cars.
   G A *city* is a unit of local government.

2. P We study and sometimes write a *poem* in our English lessons.
   G A *poem*, an essay and a story are examples of creative writing.
   M A Shakespearean sonnet is a *poem*.
   A The *poem* is a literary form.

3. G *Playing* is a universal activity among higher animals.
   A *Playing* is frequently discussed in studies of children.
   M *Playing* is tennis and football.
   P *Playing* is popular with young people.

4. A The *motorway* commends itself to many as an idea for improving motor travel.
   G The *motorway* is a civil engineering construction.
   M The *motorway* system includes M1 and M2.
   P There is a *motorway* junction near Bristol where M4 and M5 connect up.

5. M *Justice* was done when the train robbers were convicted.
   A *Justice* is a basic principle underlying the continued existence of society.
   P *Justice* is the punishment of a criminal.
   G *Justice* is a human practice making it possible for people to live together more securely.

6. A The idea of a *helper* is a human ideal which is at least as old as human written records.
   M The Good Samaritan was a *helper* of an injured traveller.
   G A *helper* is an honest man and a kindly individual.
   P A *helper* is needed for caring for the needs of old people.

7. A *Adolescent* is a term used in discussion about the needs and capabilities of young people.
   G. An *adolescent* is a growing person.
   M A teenage girl is an *adolescent.*
   P The *adolescent* is often said to be an emotionally sensitive person.

8. G A *creature without a backbone* is a living organism.
   P Many a *creature without a backbone* lives in water.
   M A snail is a *creature without a backbone.*
   A The *creature without a backbone* is thought of as a main zoological grouping.

9. M Ireland is an *island.*
   G An *island* is a feature of the earth's surface.
   P An *island* is often rocky.
   A The *island* is a geographical notion.

10. P A bicycle is an *object moving under the pull of gravity.*
    G An *object moving under the pull of gravity* is something acted upon by a force.
    A An *object moving under the pull of gravity* is an idea in physics.
    M A ball rolling down a hill is an *object moving under the pull of gravity.*

11. M Antarctica is a *Polar Region.*
    A The *Polar Region* is a geographical definition.
    G A *Polar Region* is part of the earth's surface.
    P The *Polar Region* is a home of seals.

12. G A *plane surface* is a common spatial feature of many modern buildings.
    M A billiard table top is a *plane surface.*
    A The *plane surface* is an important geometrical construct used in studying solid shapes.
    P A *plane surface* is a horizontal area.

13. G *Aggression* is an unpleasant activity carried out by people against each other.
    M Bullying at school is *aggression.*
    P *Aggression* is making someone do something.
    A *Aggression* is a psychological concept.

14. G A *team of eleven players* is a group of people.
    A A *team of eleven players* is a group idea.
    P A *team of eleven players* wins the Cup.
    M The English football side is a *team of eleven players.*

15. M Mars is a *planet.*
    P A *planet* is greatly interesting to an astronomer.
    A The *planet* is an astronomical class.
    G A *planet* is a heavenly body.

16. P The *shop* is next door to the bank.
M The greengrocer's is a *shop.*
G A *shop* is a place where things and materials are bought and sold.
A. A *shop* is a trading notion.

17. A Together with the sculptor, the *painter* forms a class of creative worker.
G A *painter* is an artistic person.
M A *painter* is a person like Picasso.
P A *painter* is somebody who likes colour.

18. A *One day and the next day* is an example of the concept of order.
P *One day and the next day* is something occurring every day.
M Tuesday and Wednesday in any one week are *one day and the next day.*
G *One day and the next day* is an example of an event followed by another.

19. A A *savings account* is a banking notion.
G A *savings account* is a way of economy for a person.
P A *savings account* is a means of getting good interest on your money.
M My building society deposit account is a *savings account.*

20. M A *plumber* is like the man who mended our burst water pipe.
P A *plumber* is a member of the Building Workers' Union.
G A *plumber* is a skilled craftsman.
A *Plumber* is one of the employment categories.

21. M General De Gaulle was a *Frenchman.*
A *Frenchman* is a national description.
P A *Frenchman* is a neighbour of the Belgians.
G A *Frenchman* is a European.

APPENDIX 2

PERCENTAGE FREQUENCIES OF RESPONSES TO ITEMS.

Item	Concept	Population L (N=218)				Population B (N=200)			
		A	G	M	P	A	G	M	P
1	City	16	24	31	28	22	23	24	32
2	Poem	17	54	9	21	15	51	11	24
3	Playing	14	6	15	65	19	7	18	57
4	Motorway	57	12	27	4	61	12	21	7
5	Justice	8	49	16	28	12	26	12	50
6	Helper	11	39	28	22	13	33	26	29
7	Adolescent	18	32	31	19	22	33	20	26
8	Creature with no backbone	30	22	28	19	38	18	24	21
9	Island	27	35	28	9	20	46	28	7
10	Object moving under gravity	23	31	31	14	15	43	32	11
11	Polar Region	27	36	26	12	25	23	32	21
12	Plane surface	22	10	40	28	23	12	32	34
13	Aggression	12	56	20	12	19	49	16	17
14	Team of eleven players	9	27	59	5	10	25	57	9
15	Planet	10	13	45	32	14	18	36	33
16	Shop	24	53	20	3	23	62	11	4
17	Painter	31	47	8	14	40	43	6	12
18	One day and the next day	12	39	24	24	14	47	31	9
19	Savings account	24	17	17	42	21	23	14	43
20	Plumber	19	31	37	12	20	34	32	15
21	Frenchman	19	45	28	8	30	38	27	6

Br. J. educ. Psychol., 45, 189-197, 1975

VERNACULAR TEST INSTRUCTIONS IN RELATION TO COGNITIVE TASK BEHAVIOUR AMONG HIGHLAND CHILDREN OF PAPUA NEW GUINEA

By M. KELLY and HUGH PHILP
(*School of Education, Macquarie University, New South Wales*)

Summary. Analysis of results of tests of cognitive development in Papua New Guinea children indicated that membership of different language/culture groups was an important source of variance. Three groups each of 48 children were selected from one language culture group and equated for age and sex. Two groups were made up of school children, equated for school experience; the third group had no school experience. One group of school children was tested in English, the language of the school; the other two groups were tested in the vernacular common to all three groups. Three tasks were used : an equivalence grouping of pictures, a class inclusion test, and conservation of length. All three tests were chosen because there were language differences between English and the vernacular in labelling the components of the task. Significant differences were found between the groups in their performance on the tasks. These differences were related both to amount of school experience and language of testing but the latter was, by far, the more significant.

INTRODUCTION

If cognitive development follows an identical course in all cultures and proceeds *parallel* with the learning and use of language, or if it influences language, then there are major implications for curriculum construction. There need be no major differences in basic structure—for example, mathematics and science may be developed and taught in much the same way in all societies. On the other hand, if language mediates development, if the child's view of the world is affected by the language he speaks, then curriculum (and methods of teaching) will have to be designed and adapted to conform.

Our own interest is explicitly in the implications of both theory and experiment for education—for what happens to children in classrooms. At the empirical level Piaget and Bruner have had most to say about the basic issues and a detailed, if somewhat biased, discussion of their positions is given in the introduction to the recent volume, edited by Anglin (Bruner, 1973) : we say ' biased ' because, like ourselves, the editor comes down broadly on the side of Sapir-Whorf and Bruner. Our view would be essentially that of Bridgman : " It begins to look as though formal logic as we know it, is an attribute of the group of Indo-European languages with certain grammatical features" (Bridgman, 1958). The data presented below are relevant to Bridgman's proposition.

The data from the " Product and Process " study, reported in Philp and Kelly (1974), showed a high degree of uniformity across cultures for task difficulty (product) and some major differences across cultures in method of solution (process) in a series of tasks based on the work of Piaget and Bruner. The data were reported in general terms for each culture group, partly because of lack of space and partly because finer analysis of the interactions between such variables as sex, sub-culture groups and age had not been completed.

The relationships among a series of these variables have now been explored for the sample of children drawn from three areas of Papua New Guinea. It was found that the most frequent source of significant variance in this sample was membership of different linguistic/cultural groups (Kelly, 1971a, 1971b, 1971c). Across tasks, this source of variance was significant more frequently than any of the other sources, such as sex, age or level of schooling. We now present data from a study designed to disentangle some of the effects of language from the effects of the material and action aspects of the culture. These (non-linguistic) social effects are referred to hereafter as ' action and material culture.' This distinction is obviously one of convenience only. We would be far from denying the importance of language in all aspects of culture—indeed, we would assert its primary value.

In order to test the hypotheses discussed above we argued—

(*a*) that children tested in their own vernacular language would perform similarly on Piaget-Bruner type tasks involving language whether or not they had been to school, and

(*b*) that children at school would perform differently when tested in the language of school instruction (English) than when tested in the vernacular.

The study involved one conservation and two classification tasks, one of which involved class inclusion. The subjects were matched groups of children, with and without school experience, tested in the language of school instruction (English) or the vernacular (Melpa), according to the design described below.

METHOD

Sample.

The subjects were children from the Melpa speaking area of the West Highlands of Papua New Guinea. In this area children who had never been to school were available, together with children from the same villages who were at school and had been in attendance for four or six years. All the children spoke Melpa as their vernacular, and the ' school children ' in addition spoke English as the language of instruction used in the school. Three groups of children were tested. Two groups were drawn from the school population and one from the non-school population. All three subsamples were matched for age and sex and, in addition, the two school samples were matched for school experience, both in terms of years of schooling and classes (i.e., same teacher). One school subsample and the non-school (village) subsample were tested in the vernacular, Melpa. The other school subsample was tested in English. Table 1 sets out the numbers in the cells for the total sample.

TABLE 1

COMPOSITION AND DESIGNATION OF SAMPLES.

	Sex	Years of Schooling			Total
		4	6	None	
Tested in English	M	* (A1) 12	(A2) 12	—	24
	F	12	12	—	24
Tested in Melpa..........	M	(B1) 12	(B2) 12	(C1) 24	48
	F	12	12	(C2) 24	48
Total		48	48	48	144

* Code numbers given to groups.

By comparing the performance of the A1 and A2 samples with B1 and B2, the effect of school experience could be held constant, while language of testing was varied. Similarly, a comparison of B1 and B2 with C1 and C2 held language of testing constant while varying school experience.

Language of Testing.

This refers to the language used in the testing situation, and includes instructions, explanations, commands and responses. If the test group was a ' vernacular ' one, all language used in the test situation was Melpa. Response behaviour was scored as well as response language. From earlier experimental evidence both languages of testing were known to be adequate for giving the necessary instructions and explanations of the tasks. Tests in the vernacular were administered by indigenous Melpa speakers, who were given an intensive training programme until reliability of instruction approached unity. Testing in English was administered by indigenous English speakers.

Tasks.

Three criterion tasks were used :

(1) *Conservation of length.* Sticks of pit-pit wood 10 cm long were equated, displaced by one-third, equated, displaced by one-third and separated by six inches and again equated. Performance was scored for correct response and for correct response with reason as in the standard Piaget procedures. An attempt was also made to analyse the nature of the 'reason' in terms of (*a*) compensation or (*b*) reversibility, as described below.

(2) *Equivalence of ' common objects.'* This task used colour photographs of five groups, each containing five stimulus objects drawn from, and common in, the Melpa culture. Earlier studies (Philp and Kelly, 1974) had indicated that there was no difference in classification behaviour of Melpa children whether real objects or colour photographs were used.

The five groups of objects varied in the possession of a group name (single word or phrase) over the two languages of testing. The variation is set out in Table 2. It is stressed that the variation is in the name for the *group* as a group. All the *individual* objects had names in both languages.

TABLE 2

BREAKDOWN OF GROUPS OF PHOTOGRAPHS ACCORDING TO LANGUAGE LABEL.

English Groupings	Melpa		English	
	Word name	Phrase name	Word name	Phrase name
1. Household goods	No	No	No	Yes
2. Clothes	No	Yes	Yes	Yes
3. Tools	No	Yes	Yes	Yes
4. Food plants	No	Yes	Yes	Yes
5. Trees	Yes	No	Yes	Yes

Four of these five groups possess group label words commonly used in the vernacular. The fifth group, household goods, contains *objects* common in the language and common in the culture, but with no label as a *group* in Melpa.

Four sets of Melpa ' label ' groups were chosen to explore the effect of combination of Melpa label groups with the Melpa ' no label ' groups. Children were asked to " sort out the picture into five piles." Any group of five photographs missed was assembled and the child asked " Can those things go together ? " In both cases the child was asked how, or why, the members of any agreed group went together. Finally, the following combinations of groups of five were assembled : household goods and tools ; household goods and clothes ; household goods and tools and clothes ; and food-plants and trees. For each combination the children were asked " Can those things go together ? " The how/why question followed if they accepted the grouping.

(3) *Hierarchy or class inclusion.* This task was administered in two ways, using a split half design. One method concerned the ability of the children progressively to divide the constructed hierarchy according to the classification principles on which it was made up. The second method explored whether they were aware of the inclusive/exclusive and subsumption dimensions involved in Piaget's test of class inclusion. The hierarchy was made up of coloured photographs of common plants in the Melpa culture as set out in Table 3. Compared with English, Melpa possesses very few words to express hierarchical superordination and we were unable to find a hierarchy with more than the three levels of the task.

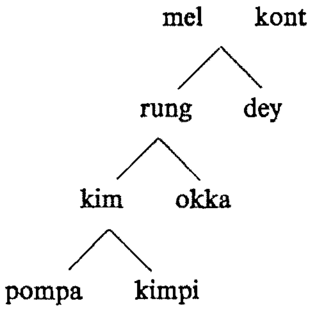
TABLE 3

CONTENT OF THE LEVELS OF CLASS INCLUSION USED IN THE HIERARCHY.

Level		Superordinate name for group
1	4 Kimpi+2 Pompa (4 vegetable A+2 vegetable B)	6 Kim (6 leaf vegetable)
2	6 Kim+2 Okka (6 leaf vegetable+2 Kau Kau)*	8 Rung (8 foods)
3	8 Rung+2 Dey (8 Foods+2 trees)	10 Mel Kont (10 living things)

*Kau Kau is a root vegetable of the sweet potato family : there are many varieties, each with its own name.

The hierarchy can be represented diagrammatically as set out below :



In Method 'A', the children were asked to divide the 10 living things (mel kont) into two piles, the larger of these (rung) into two, and so on. Scoring was based on adherence to the scheme as outlined above. Note was taken of the composition of any division used other than that decided *a priori* and outlined here. These ' deviant ' divisions are not reported. At all points the children were asked to identify the pictures before division took place. In Method ' B ' the child was asked to identify the array at level 1, then asked if these were all ' kim.' On agreement, he was then asked "Are there more pompa than kimpi ? Are there more kimpi than pompa ? Are there more pompa than kim ? Are there more kimpi than kim ? " This procedure was followed with levels 2 and 3. Scoring was based on successful answers to the inclusion question after correct identification of the array.

RESULTS

(1) *Conservation of Length.*

It will be remembered that the experimental design is such that the crucial comparisons between pairs of subsamples are two—school English *versus* school vernacular and school vernacular *versus* village vernacular. Table 4 presents these comparisons for the task involving conservation of length.

TABLE 4

PERCENTAGE OF SUBSAMPLE SUCCEEDING IN CONSERVATION OF LENGTH.

Response	Language of testing		
	School English (N=48)	School Vernacular (N=48)	Village Vernacular (N=48)
Correct—with or without reason	45·8	37·3	20·8
Correct—with reason	37·5	27·0	10·4

The difference in the percentages of the subsamples succeeding between ' school ' children tested in Melpa is not statistically significant at the ·05 level. The second comparison, between ' school ' children and ' village ' children, both tested in the vernacular is, however, highly significant ($P < \cdot 01$). That is, schooling makes a difference to the capacity to conserve length.

Since we are mainly concerned in this article with language, it is of particular interest to note that the proportions of children able to ' give a reason ' for conservation related to the total able to succeed in the task ' with or without ' reason decreases significantly from ' School English ' (·82) to ' School Vernacular ' (·72) to ' Village ' (·50).

We point out—without claiming a causal relationship—that it is easy, in the Melpa language, to express a reason for conservation based on compensation (for example, " the bit sticking out at that end is the same as the bit at the other end "). However, it is very difficult, although not impossible, to express reversibility because there is no conditional mood (for example, one cannot directly say " if you pushed them together again, they would fit exactly "). No child from school or village used the possible circumlocution in the vernacular by giving as a reason for conservation an explanation based on reversibility.

(2) *Equivalence of ' common objects.'*

Differences between groups for the task involving household goods when the school English and school vernacular used are compared in Table 5. Differences for household goods and clothes are in the same direction, but are not statistically significant. No significant differences were found between school vernacular and village vernacular groups in the tasks involving household goods (alone, plus tools, plus clothes, plus tools and clothes) or trees plus food plants.

TABLE 5

PERCENTAGE OF SUBSAMPLE SUCCEEDING IN GROUPING THE PICTURES.

Group	School English	School Vernacular	Significance P<
Household goods	52·1	20·9	·001
Household goods, plus tools	39·6	20·9	·02
Household goods, plus tools, plus clothes....	37·5	23·0	·05

(3) *Hierarchy or class inclusion.*

Method 'A'.

Percentages successful in dividing the photographs according to design criteria are given in Table 6. Clearly school experience has a highly significant effect on ability to divide classes of objects, *even when they are made up from the vernacular.* This supports the contention that, even where the language is perfectly adequate to form a hierarchy, the children *do not*, in fact, do this as a matter of course.

TABLE 6

PERCENTAGE OF SUBSAMPLE SUCCEEDING IN CLASSIFICATION OF HIERARCHY (METHOD A).

	School English	School Vernacular	Village Vernacular
Percentage	50·0	45·8	16·6

Method ' B.'

Table 7 presents the percentage of each subsample achieving success at each level. The percentages are gross, that is all successful subjects are included for each level, irrespective of success or failure of the individual at an earlier level.

TABLE 7

PERCENTAGE OF SUBSAMPLE SUCCEEDING IN CLASSIFICATION OF HIERARCHY (METHOD B).

Success at	School English	School Vernacular	Village Vernacular
' kim ' level	81·2	70·8	60·5
' rung ' level	43.7	22·9	31·2
' mel kont ' level	58·3	52·0	41·6

The only comparison between criterion subsamples which reaches statistical significance ($P < \cdot 01$) lies between school English and school vernacular at the 'rung' inclusion point, though the general trend is school English>school vernacular>village vernacular. It should be noted that the words 'kim' and 'mel kont,' like the English equivalents 'leaf vegetable' and 'living things,' are terms which have a clear inclusion/exclusion denotation. The term 'rung,' which is normally translated as 'food,' unlike its English 'equivalent,' is not as precise in demarcation since it is more properly thought of in common usage as 'prepared for eating' or 'cooked.' However, it is also used in the sense of 'can be prepared for eating' and in this sense it is a reasonable parallel to the English 'food.' We cannot tell, from our data, whether children thought of *rung* in this inclusive sense or not.

The effects of this distinction become clearer if the tasks are considered as forming a very simple Guttman scale. That is, on the grounds of the logic of English, any child able to classify objects as 'mel kont' would also be expected to succeed in classifying them as 'rung' and, similarly, success at the 'rung' level would imply success with 'kim.' With a three-item scale, there are eight scale types, of which four are 'perfect.' Table 8 presents an analysis of children's responses to each of the three levels in the hierarchy following the usual convention for success (+) and failure (—).

TABLE 8

ANALYSIS OF RESPONSES OF SUBSAMPLES BY SCALE TYPES DERIVED FROM THREE LEVELS OF THE HIERARCHY.

	School English			School Vernacular			Village* Vernacular			Total		
	4	6	T	4	6	T	4	6	T	4	6	T
'Perfect' Types												
+ + +	3	11	14	3	4	7	2	3	5	8	18	26
— + +	2	2	4	0	2	2	2	1	3	4	5	9
— — +	7	3	10	7	8	15	10	3	13	24	14	38
— — —	3	1	4	5	1	6	2	6	8	10	8	18
Total	15	17	32	15	15	30	16	13	29	46	45	91
'Non-Perfect' Types												
+ + —	0	1	1	0	4	4	2	1	3	2	6	8
+ — —	0	2	2	3	1	4	1	3	4	4	6	10
+ — +	8	3	11	6	4	10	1	7	8	15	14	29
— + —	1	1	2	0	0	0	4	0	4	5	1	6
Total	9	7	16	9	9	18	8	11	19	26	27	53

* Subscripts 4 and 6 refer to groups of children equivalent in age to the children with four and six years of schooling.

The coefficient of reproducibility for the total group is 0·88, which indicates that, in strict statistical terms, the scale is not acceptable—the logic and data do not agree.

However, from our point of view, it is of considerable interest that by far the greatest number of non-scale responses fall in the + — + category. That is, 29 of the 144 children were able to classify *all* the stimuli as 'mel kont,' and were also able to classify the six leaf vegetables as 'kim,' *but* did not combine the ' kim ' and ' okka ' as ' rung '—which is what might have been expected from the linguistic considerations discussed above. The ' logic of Melpa ' is not necessarily the ' logic of English.' Moreover, as is apparent from Table 8, there are few differences among the three groups, *except* in the greater tendency of the ' School English ' group, to form the ' mel kont ' classification. As a matter of interest, the scales for the four years of schooling and six years of schooling have almost identical indices of reproducibility (·86 and ·90) and, on inspection, there is an evident developmental trend in all three groups.

Definitive conclusions cannot, of course, be drawn from a scale as simple as this, but at least it points in the predicted direction. The language tends to be reflected in the logic.

DISCUSSION

Conservation.

There appears to have been both a ' language ' and a material, or action, ' culture ' effect in these results, though statistical significance was achieved only from the ' culture ' effect. The effect of the school/cultural difference is interesting in the light of Piaget's often expressed opinion that the structures he describes are not taught in school (Piaget, 1968). The absolute proportion of children achieving success drops when ability to explain the answer is taken into account. However, it drops in all three groups, with the largest proportionate drop being in the groups of village children tested in the vernacular. This may reflect the fact that it is rare for village children to be required to explain their actions, while school children are often asked for reasons. Moreover, some kinds of explanations cannot be expressed in Melpa without a great deal of circumlocution which children in our sample did not attempt to use.

Equivalence.

Since the subsamples were matched for age and sex and the school subsamples matched for school experience in addition, it can only be concluded that the use of a language (for instruction in the task) which lacked words to name the group of household goods in some way inhibited the behaviour of the school vernacular subsample. It is assumed that both the school groups could be considered equally able in English and in superordination. With groups of photographs where this disability did not exist the behaviour of all three subsamples was so similar as to be nearly identical.

Hierarchy.

It would appear that there is an effect of both school experience and language of testing on the ability to form hierarchies of classification as well as in superordination *per se*. The effect of the language of testing may have been due to the more precise inclusion/exclusion denotation of the English term ' food ' compared to the Melpa term ' rung,' even when both terms were equally well-known by the children.

Implications.

The pattern of the ' hierarchy ' task fitted the language and culture of the subjects. However, left to themselves, the children did not automatically use this pattern. This suggests that Melpa children are not conscious of the patterns of classification of their own language and that it would probably be useful to attempt to make them aware of these before imposing foreign and alien patterns such as are found in the mathematics and science concepts of Western technology. Domino and card games could possibly help to achieve this as early as the first school grade.

The difference reported in the ' common objects ' task indicate that verbal labels are important props to classification behaviour. Perhaps Ausubel is right in emphasising the need for *meaningful* verbal learning. As yet we know very little about ' translation effects ' across language and how they affect this meaning. The present results are similar to the findings of Kelly, Tenezakis and Huntsman (1973) and Keats and Keats (1972), though children in both these studies possessed equivalent terms in both their languages. It points to the importance of the choice of language of instruction in the school both for adequacy of vocabulary and ' generalisability ' to life outside the school where a different language is used. It may be that here, too, a great awareness of the structures of the vernacular language and culture can assist generalisation from the school life to the community life.

On the more theoretical level, the data tend in the direction of Bruner, rather than Piaget, that is, they indicate that language mediates thought and logic, rather than being the servant of thought and the mirror of logic. We would agree with Piaget that on one level the issues are epistemological, but assert that to a considerable degree the child's view of the world is mediated, perhaps even to some extent determined by, the language he speaks. In the equivalence task for example, the children may have mentally rejected classifications after the first ' justification ' had been required, because they lacked a ' handy ' word. Yet the difference lies between the two groups of school children who, presumably, were equally able in English where a ' handy ' word did exist. It would seem then that the language used by the tester in some way ' directed ' the child to behave in one way rather than another, though he was presumably equally able to behave in either fashion. Thus, we have argued that language is something more than the slave of thought. Finally, and somewhat more tentatively, children may be able to operate effectively in terms of more than one epistemology, depending on the language being used.

REFERENCES

BRIDGMAN, P. W. (1958). Quo Vadis. *Daedalus*, 87, 85-93.

BRUNER, J. S. (1973). *Beyond the Information Given*. New York : Norton.

HOIJER, H. (1954). The Sapir-Whorf Hypothesis. In HOIJER, H. (Ed.), *Language in Culture*. Chicago : University of Chicago Press.

KEATS, J., and KEATS, D. (1972). The effect of language on concept acquisition in bilingual children. Paper read to the Australian Association for Research in Education Conference, Canberra, November, 1972.

KELLY, M. (1971a). Some aspects of conservation of quantity and length in Papua/New Guinea in relation to language, sex and years of school. *Papua New Guinea J. Educ.*, 7, 55-60.

KELLY, M. (1971b). A two criteria classification matrix with some Papua New Guinea Children. *Papua New Guinea J. Educ.*, 7, 46-49.

KELLY, M. (1971c). The validity of Bruner's concept of modes of representation of reality with a sample of Papua New Guinea children. *Papua New Guinea J. Educ.*, 7, 33-37.

KELLY, M., TENEZAKIS, M., and HUNTSMAN, R. (1973). A report of some unusual conservation behaviour in children exposed to two cultures. *Br. J. educ. Psychol.*, 43, 181-182.

LURIA, A. R. (1959). The directive function of speech in development and dissolution. *Word*, 15, 341-352.

PHILP, H. (1973). Mathematical education in developing countries—some problems of teaching and learning. In HOWSON, A. G. (Ed.), *Developments in Mathematical Education*. Cambridge University Press.

PHILP, H., and KELLY, M. (1974). Product and process in cognitive development : some comparative data on the performance of school age children in different cultures. *Br. J. educ. Psychol.*, 44, 248-265.

PIAGET, J. (1968). La Naissance de l'intelligence. *L'Express*, 23-29 December, 1968, 47-54.

SINCLAIR-DE ZWART, H. (1969). Development Psycholinguistics. In ELKIND, D., and FLAVELL, J. (Eds.), *Studies in Cognitive Development*. Oxford : Oxford University Press.

WHORF, B. L. (1956). *Language, Thought and Reality*. Cambridge, Mass. : Technology Press.

(*Manuscript received 25th July, 1974*)

Br. J. educ. Psychol., 45, 198-205, 1975

PERFORMANCE ON TWO REASONING TASKS IN RELATION TO INTELLIGENCE, DIVERGENCE AND INTERFERENCE PRONENESS

BY ELIZABETH R. VALENTINE
(*Department of Psychology, Bedford College, University of London*)

SUMMARY. In this study performance on reasoning tasks commonly used in psychological experiments was compared with scores on tests of convergent and divergent thinking. The relations between performance on a deductive and an inductive task, and scores on the AH5, Uses of Objects and Stroop tests were investigated in 38 first-year undergraduates. The results showed a positive correlation between performance on the deductive task and AH5 scores, a negative correlation between performance on the inductive task and scores on the Uses of Objects test, and a negative correlation between AH5 and Stroop scores. Implications for an analysis of the two reasoning tasks were discussed.

INTRODUCTION

THE aim of the present study was to relate performance on two reasoning tasks extensively studied by experimental psychologists (Wason and Johnson-Laird, 1972), to that on three mental tests, namely, the AH5, Uses of Objects and the Stroop test. The latter were selected partly on account of their importance in differential psychology and hence the amount of information already accumulated about them, and partly on the basis of what is known about the nature of the processes involved in the reasoning problems.

The first of these reasoning tasks, which we may call the ' Four Card Problem ' following Lunzer *et al.* (1972), is a deductive task examined in a number of experiments by Wason (1964, 1968, *et seq.*). It involves making inferences from a conditional statement of the form : ' If p then q.' In the version used here the subject is presented with four cards showing respectively, a vowel, a consonant, an even number and an odd number. He is told that every card has a letter on one side and a number on the other. He is then presented with a sentence such as " If there is a vowel on one side of any card, then there is an even number on its other side " and asked to say which cards he needs to turn over in order to find out whether the statement is true or false. For the present example we may symbolise a vowel as ' p,' a consonant as ' not-p,' an even number as ' q ' and an odd number as ' not-q.' The correct answer is to select ' p ' and ' not-q,' as either an odd number on the other side of the vowel or a vowel on the other side of the odd number would falsify the statement. The majority of subjects select ' p ' and ' q ' or just ' p.' Hence, two errors are commonly committed :

(1) *Failure to select 'not-q.'* There may be two reasons for this: (*a*) lack of reversibility, that is, a failure to accept that ' not-q ' on the visible side and ' p ' on the other side of the same card is the same as ' p ' on the visible side and ' not-q ' on the other side of a card ; and (*b*) failure to see that a card should be selected if it could potentially falsify the statement.

(2) *The selection of ' q.'* The error here lies in selecting a card that could not falsify the statement. Johnson-Laird and Wason (1970) have proposed an information processing model to account for performance on this task, which postulates three degrees of insight in terms of increasing recognition of the

significance of falsification. Evans and Lynch (1973) have presented evidence which suggests that subjects' behaviour is determined by ' matching bias ' (a tendency to respond with values which match those mentioned in the rule) rather than by a preference for verification over falsification strategies.

As the Four Card Problem is a deductive task, it was hypothesised that one factor likely to lead to success was convergent intelligence. Hence, it was predicted that performance on this task would be positively correlated with scores on the AH5. This would also be consistent with Piaget's stress on reversibility as a fundamental feature of intelligence and the suggestion that failure on the task may be associated with lack of reversibility. There are no specific tests of ability to perceive the significance of falsification or avoidance of ' matching bias ' but, in the event of the predicted correlation obtaining, it would be tempting to consider these as aspects of general intelligence. The ' matching bias ' hypothesis would also be consistent with Piaget's analysis of intelligence which demands the relinquishing of perceptually dominated, in favour of abstract, modes of thinking.

The second problem, which we may call the rule problem (hereafter referred to as the ' Rule '), models some aspects of scientific thinking and was first examined in an experiment by Wason (1960). The subject is presented with the numbers 2, 4 and 6 and told that they conform to a simple relational rule which the experimenter has in mind. The task is to discover the rule by generating other sets of three numbers, after each of which he is told whether or not they conform to the rule. He is instructed to announce the rule when and only when he is confident about it.

The rule is, in fact, "Any three numbers in increasing order of magnitude," purposely designed to encourage sufficient but not necessary hypotheses, for example, ' increasing even numbers.' Subjects show a marked reluctance to attempt falsification of their hypotheses, the majority of them making one or more incorrect announcements. Those who were correct on their first attempt showed a significantly greater tendency to vary their hypotheses.

It was hypothesised that divergence would be a factor likely to lead to success on the Rule, on the grounds that divergent students would generate a greater number and variety of hypotheses and hence be more likely to hit on the rule. It was thus predicted that there would be a positive correlation between performance on the Rule and Uses of Objects score. As one of the main stumbling blocks in the task is inability to relinquish obvious hypotheses, it was postulated that another factor involved might be the inability to inhibit a dominant response in favour of a less dominant conflicting response. The Stroop test has been considered as measuring such interference proneness (Jaensch *et al.*, 1929 ; Stroop, 1935). Hence, it was predicted that success on the Rule would be negatively correlated with interference proneness scores.

METHOD

Subjects.

The sample consisted of 11 male and 27 female first-year psychology students at Bedford College, London.

Tests and procedure.

Five tests were administered. Details of the procedure for each were as follows :

(1) *The Four Card Problem* was presented on a sheet of paper with the

following instructions :

> " This is a short reasoning task ; it is not an intelligence test. Below are represented four cards. You are given the information that each card has a letter on one side and a number on the other.
>
> (The four cards were represented here)
>
> You are then given the following sentence:
>
> (E.g.) " If there is a consonant on one side of any card then there is an even number on its other side."
>
> Your task is to write down below which of the cards you need to turn over in order to find out whether this statement is true or false."

No time limit was imposed ; all students completed the task within a few minutes. The four cards were represented by four squares containing a vowel, a consonant, an even number and an odd number, respectively. All possible non-confusable letters and single digit numbers were used ; they were selected at random and all possible orderings of the four types occurred as equally often as possible.

(2) The instructions for the *Rule* were as follows :

> " You will be given three numbers which conform to a simple rule that I have in mind. This rule is concerned with the relation between any three numbers and not with their absolute magnitude, i.e., it is not a rule like all numbers above (or below) 50, etc.
>
> Your aim is to discover this rule by writing down sets of three numbers, together with reasons for your choice of them. After you have written down each set, I shall tell you whether your numbers conform to the rule or not, and you can make a note of this outcome on the record sheet provided. There is no time limit but you should try to discover this rule by citing the minimum sets of numbers. Remember that your aim is not simply to find the numbers which conform to the rule, but to discover the rule itself. When you feel highly confident that you have discovered it, *and not before*, you are to write it down and tell me what it is. Have you any questions ? "

The students were presented with a sheet of paper on which were four columns. In the first the numbers 2, 4, 6 were shown as an example ; the second column was headed ' reasons for choice ' ; the third and fourth were headed, respectively, ' the rule holds for these numbers ' and ' the rule does not hold for these numbers,' so that the outcome of each trial could be recorded. After solving the problem each student was asked whether he considered the answer fair (all said they did) and requested not to discuss the experiment with anyone.

(3) The conventional *intelligence test* used was the AH5 revised version (Heim, 1968). The instructions in the manual were followed.

(4) The *divergence test* employed was Uses of Objects (Getzels and Jackson, 1962 ; Hudson, 1966). The students were given a sheet of paper, on which were the following instructions :

> " This is a short test of your creative imagination. I want you to think up as many uses as you can for the objects below. E.g., if the object was a book, you might say that it could be used as a cricket bat, a weapon, a door-step, to light a fire and so on. Below are five everyday objects. Think of as many different uses as you can for each and write them down on the page below."

The objects were a barrel, a paper clip, a tin of boot polish, a brick and a blanket. A total of 15 minutes was allowed, the time to be distributed between objects as desired.

(5) The procedure for the *Stroop test* was based on that of Gardner *et al.* (1959). The material for the experimental condition consisted of 20 rows of five colour words (red, blue, green, yellow and black) written in one of these five colours with the following constraints : the colour in which the word was written was never the same as its name and no two successive colours or colour names were the same. The material for the control condition consisted of rows of coloured asterisks, whose colours and positions corresponded exactly to the letters in the experimental condition. After a practice row in each case, each student was timed for one trial under the control condition followed by one trial under the experimental condition. The instructions were as follows :

> *Control condition* : " I'm going to present you with a page of blocks of asterisks of different colours. I want you to go along each row, starting at the top left, calling out the colours as fast as you can but without making any mistakes ; and be careful to keep track of the rows so as not to miss or repeat any. To check you've the idea, try this row for practice . . . O.K. ? Go."
>
> *Experimental condition* :. " This time I'm going to present you with a page of colour names written in different colours, e.g., the word ' red ' might be written in blue ink. (The colours will be the same as for the asterisks.) I want you to ignore the colour names and go along the rows as before calling out the colours in which they are written as fast as you can but without making any mistakes. Do you understand ? Try this row for practice . . . O.K. ? Remember, ignore the names, just call out the colours. Go."

The five tests were administered as follows : The subjects received the intelligence test in groups when they attended for interview for selection for the course. However, it was only possible to obtain data in this way for 31 of them. The Four Card Problem, followed by the divergence test, was presented to the students in three separate groups for administrative purposes. The Stroop test and the Rule were presented individually. Nine of the subjects had to be rejected from the Rule as they were already familiar with it. In addition, one student withdrew from the course between testing sessions and another had to be rejected from the two reasoning tasks, as it was clear that he did not fully understand the instructions.

RESULTS

The Four Card Problem.

In general the pattern of responses obtained was similar to those of other studies. Only five out of 37 students made the correct selection, ' p ' and ' not-q.' The most frequent response pattern was ' p ' and ' q,' and the next most frequent was ' p.' However, there were several bizarre responses, for example, selections including ' not-p,' and no cases of a selection of ' p,' ' q ' and ' not-q,' which would have been indicative of partial insight as defined by Johnson-Laird and Wason (1970) ; hence, degrees of insight could not be used as a basis for scoring. Instead, a score of +1 was given for each correct selection or omission, and a score of −1 for each incorrect selection or omission.

The Rule.

All the students attained the rule eventually but the majority made at least one incorrect announcement. Nine were correct on their first announcement, seven on their second, five on their third, and six on their fourth. None required more than four attempts.

AH5.

The standard scoring procedure was adopted, giving a grade from A to E for the total test, and also separate scores for Part I (verbal and numerical) and Part II (diagrammatic). A represents the top 10 per cent, B the next 20 per cent, C the middle 40 per cent and D the next 20 per cent. No student scored E, representing the bottom 10 per cent.

Uses of Objects.

A fairly generous scoring procedure was adopted, each separate use suggested for an object being counted as one, unless there were obvious repetitions. Scores for the five objects were totalled. The distribution was slightly skewed, ranging from 16 to 77, with a mode and median of 35 and a mean of 37·5.

Stroop.

Times taken to read out the colours were recorded for the control and experimental conditions. Errors were rare but were penalised by adding a correction term to the reading time. This term was calculated by multiplying the reading time per unit (i.e., the reading time divided by the number of items read, whether correct or incorrect) by the number of errors.

Since a highly significant regression of reading time for the experimental condition on reading time for the control condition was obtained ($r = \cdot 67$, $P < 0\cdot 001$), interference scores were computed for each student by means of the following formula : $(RT_o - RT_p)/RT_p$, where RT_o was the observed reading time and RT_p the reading time for the experimental condition predicted on the basis of the regression. (Inspection of the data suggested that the appropriate formula for the regression coefficient, b, was $(\Sigma y/x)/n$, recommended by Snedecor (1956) for data where the regression line is assumed to pass through the origin and where the variance increases as a constant function of x. This last fact also dictated the use of a ratio for the interference score.)

Statistical analysis.

The Four Card Problem.

The relation between scores on the Four Card Problem and grades on the AH5 is shown in Table 1.

TABLE 1

THE RELATION BETWEEN SCORE ON THE FOUR CARD PROBLEM AND AH5 GRADE.

		AH5 grade				
		A	B	C	D	Total
Score	5	2	1	0	1	4
on the	4	1	4	3	0	8
four	3	0	5	5	5	15
card	2	0	0	2	0	2
problem	1	0	0	1	0	1
Total		3	10	11	6	30

Kendall's S revealed a significant positive correlation (S=114, z=2·306, P=0·01, one-tailed test). Examining grades on the two separate parts of the AH5, the correlation with Part I (verbal and numerical) just fails to reach significance (S=75, z=1·532, P=0·06, one-tailed test) but that with Part II (diagrammatic) is significant (S=113, z= 2·283, P=0·01, one-tailed test).

There were no significant relations between performance on the Four Card Problem and that on either the Uses of Objects or the Stroop test.

The Rule.

The mean scores on the Uses of Objects test for students making each of the four numbers of announcements on the rule were as follows : 1 : 29·5 ; 2 : 41·6 ; 3 : 44·8 ; 4 : 41·8. Thus, an inverse relation was obtained between performance on the Rule and scores on the divergence test—there was a significant tendency for students scoring highly on the Uses of Objects test to make more announcements (S=109, z=2·351, P=0·02, two-tailed test). It might be thought that they were more impulsive. However, further analysis showed that there was a significant positive relation between score on the Uses of Objects test and number of instances produced in the Rule (S=115, z=2·396, P=0·008), but not between Uses of Objects score and number of instances generated per rule announcement. Hence, ' divergent ' students do not make announcements more frequently ; it is simply that they make more (incorrect) announcements.

There was no significant relation between number of announcements on the Rule with either AH5 grade or interference proneness score on the Stroop test.

An examination of the relation between score on the Four Card Problem and number of rules announced gave a value for Kendall's S of zero ; hence, the two problems apparently tap different abilities.

An examination of the relation between interference proneness scores on the Stroop test and AH5 grades revealed a significant negative relation (S=130, z=2·55, P=0·01, two-tailed test), such that ability to inhibit conflicting responses in the Stroop test was positively correlated with AH5 grade. This is an interesting result in view of the rather low correlations obtained by previous authors.

There was no relation between scores on the Uses of Objects and Stroop tests.

DISCUSSION

The results of the correlations of performance on the Four Card Problem with that on the other tests used show that success on this deductive task is positively correlated with convergent intelligence as measured by the AH5. The suggestion that failure to select ' not-q ' may be due to lack of reversibility is interesting in view of Piaget's stress on reversibility as a fundamental feature of intelligence. Similarly, the ' matching bias ' hypothesis of subjects' errors on the task is consonant with the Piagetian view that perceptual modes of behaviour are more primitive than abstract ones. However, the results also show that convergent intelligence is a necessary but not a sufficient condition for success on the task. This is consistent with the informal observation that several highly intelligent people, including a professor of logic, have failed to make a correct selection. Further research is required to determine what factors additional to convergent intelligence are necessary for success. No significant relation was obtained between performance on this task and scores on either the Uses of Objects or the Stroop test.

Contrary to expectation, success on the Rule problem was found to be negatively correlated with divergence as measured by the Uses of Objects test. High divergers produced significantly more announcements. They produced more instances, though not more instances per rule. Their productivity failed to lead to success on this task as they were apparently unable to subject their hypotheses to critical analysis. However, this result is consistent with the findings of several other authors (see, for example, Taylor and Williams, 1966). In the classic study by Wallach and Kogan (1965), the authors began by defining creativity in terms of associative fluency but were forced to acknowledge the addition of a critical, evaluative aspect. Similarly, Shouksmith (1970) concludes that creative problem solving demands the ability to switch from an initial, open approach to a more perseverative phase in which ideas are subjected to critical analysis. He also found that females were less likely to combine convergent intelligence with creativity and frequently showed a fluency which did not lead on to flexible problem solving. This may explain the marked inability of the present sample to combine these two forms of activity, as the majority of the students were female. It must be concluded that performance on the Rule is a better index of creative problem solving than is a score on a test of divergence. Neither Stroop interference proneness scores nor AH5 scores were found to be related to the number of rule announcements made.

The extent of the positive correlation between lack of susceptibility to interference on the Stroop test and AH5 grade is also of interest in view of the rather low correlations reported by previous investigators using other intelligence tests. Callaway (1959), Leedy (1963, quoted by Jensen and Rohwer, 1966) and Jensen (1965) all obtained correlations near zero between Stroop interference scores and Raven's Progressive Matrices ; Uhlmann (1962, also quoted by Jensen and Rohwer, *op cit.*) obtained a correlation of $r = \cdot 34$ ($P < 0 \cdot 01$) between the Stroop interference factor and the Wonderlic Personnel Test, a general intelligence test used in business and industrial situations. It is conceivable that the AH5 is sufficiently sensitive for the intelligence range tested in the present study to bring out the effect. Another possibility is that some of these investigators may have failed to allow for the increase in variance likely to be associated with slower reading times in the Stroop test and hence have contaminated their measures of interference with purely speed factors (see Results section). This is certainly true in the case of Callaway, who used a simple subtraction score.

Further research is required to determine the validity of this last finding and to explore further factors involved in these reasoning tasks. The results reported in the present study suggest that such extensions would be fruitful.

ACKNOWLEDGMENTS—I am grateful to Richard Kirby for discussion of the original ideas, Jim Kyle, Sara Meadows and Margaret Riseborough for the administration of the AH5, Kate Loewenthal for assistance with the administration of the Four Card Problem and Uses of Objects test, and John Valentine for statistical advice.

REFERENCES

CALLAWAY, E. (1959). The influence of amobarbital (amylobarbital and metamphetamine) on the focus of attention. *J. Ment. Sci.*, 105, 382-392.

EVANS, J. ST. B. T., and LYNCH, J. S. (1973). Matching bias in the selection task. *Br. J. Psychol.*, 64, 391-397.

GARDNER, R., HOLZMAN, P. S., KLEIN, G., LINTON, H., and SPENCE, D. P. (1959). Cognitive control : a study of individual consistencies in cognitive behaviour. *Psychol. Issues*, 1, 4, 1-185.

GETZELS, J. W., and JACKSON, P. W. (1962). *Creativity and Intelligence*. New York : Wiley.

HEIM, A. W. (1968). *AH5 Group Test of High-Grade Intelligence*. (Revised edit.). Windsor : National Foundation for Educational Research.

HUDSON, L. (1966). *Contrary Imaginations*. London : Methuen.

HUDSON, L. (1968). *Frames of Mind*. London : Methuen.

JAENSCH, E. R., *et al.* (1929). *Grundformen Menschlichen Seins*. Berlin : Otto Elsner.

JENSEN, A. R. (1965). Individual differences in learning : interference factor. Final Report, Co-operative Research Project No. 1867, US Office of Education.

JENSEN, A. R., and ROHWER, W. R. (1966). The Stroop colour-word test : a review. *Acta Psychol.*, 25, 36-93.

JOHNSON-LAIRD, P. N., and WASON, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychol.*, 1, 134-148.

LEEDY, H. B. (1963). An investigation of the range of cue utilisation, as a function of induced and intrinsic drive. Unpublished doctoral dissertation, Purdue University.

LUNZER, E. A., HARRISON, C., and DAVEY, M. (1972). The four-card problem and the generality of formal reasoning. *Q. Jl. exp. Psychol.* 24. 326-339.

SHOUKSMITH, G. (1970). *Intelligence, Creativity and Cognitive Style*. London : Batsford.

SNEDECOR, G. W. (1956). *Statistical Methods*. Iowa : Iowa State Univ. Press.

STROOP, J. R. (1935). Studies in interference in serial verbal reactions. *J. exp. Psychol.*, 18, 643-661.

TAYLOR, C. W., and WILLIAMS, F. E. (Eds.) (1966). *Instructional Media and Creativity*. New York . Wiley.

WALLACH, M. A., and KOGAN, N. (1965). *Modes of Thinking in Young Children*. New York : Holt, Rinehart and Winston.

WASON, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Q. Jl. exp Psychol.*, 12, 129-140.

WASON, P. C. (1964). The effect of self-contradiction on fallacious reasoning. *Q. Jl. exp. Psychol.*, 16, 30-34.

WASON, P. C. (1968). Reasoning about a rule. *Q. Jl. exp. Psychol.*, 20, 273-281.

WASON, P. C., and JOHNSON-LAIRD, P. N. (1972). *Psychology of Reasoning*. London : Batsford.

(*Manuscript received 4th October*, 1974)

Br. J. educ. Psychol., 45, 206-213, 1975

GUESSING THE ANSWER ON OBJECTIVE TESTS

BY B. CHOPPIN

(*National Foundation for Educational Research, Slough*)

SUMMARY. Using data obtained as part of the cross-cultural IEA study on academic standards, the tendency of particular pupils to guess on multiple-choice tests was measured using an index proposed by Ziller. The amount of guessing was found to increase with the amount of reading material each item contained. English school pupils appear to be more inclined to guess than are pupils in most other Western countries. Guessing is more frequent in the public sector of education than in the independent sector. Among 14-year-olds, boys were slightly, but consistently, more likely to guess than girls. This sex difference was not apparent among children aged 10.

INTRODUCTION

THE use of objective tests of attainment, and particularly the multiple-choice item format, increases year by year in our school system. Whatever psychometric advantages such tests exhibit in a research context, there is no doubt that they have been subject to criticism from educators since they were first introduced.

The feature of a multiple-choice test item that arouses the greatest hostility among teachers and examiners is that it requires the recognition of, rather than the construction of, the correct response. This has implications for both the validity and the reliability of the test ; the validity, because the test constructor is usually more interested in assessing the students' ability to solve a problem than his capacity to recognise the solution when it is shown to him ; the reliability, because the student has a very real chance (a fifth, a quarter, or sometimes even a half) of guessing the right answer, even when he has absolutely no idea how to solve the problem. Ingenious test writers are constantly finding new methods of minimising the effects of the first difficulty. They do this by inventing question structures which play down the recognition component in the solution. The best strategy to deal with the second difficulty would seem to lie in reducing the possibility of a correct guess by offering a wide number of alternative answers. This, however, has usually been found to be both inconvenient and inefficient, and as a result, a certain amount of guessing has come to be tacitly accepted as an inescapable part of a pupil's repertoire of behaviour when confronted with a multiple-choice test.

Although, as two recent reviews have shown (Choppin, 1975 ; Diamond and Evans, 1973), there has been a great amount of generally rather unsuccessful research into methods of minimising guessing, or the effects of guessing, there has been relatively little study of guessing on tests as a behaviour in its own right. This seems strange, since it is a matter of common experience that there are good guessers and bad guessers, confident guessers and nervous ones.

Correcting for Guessing.

It has long been recognised that individual variations in guessing could lead to a contamination of achievement test scores. For example, let us consider two students of equal ability taking a multiple-choice geography test. One student is confident and the other very nervous. The nervous student solves the problems he knows how to solve, and leaves the answers to the rest of the questions blank. The confident student not merely responds correctly to the questions he knows but also hazards a guess at all the others. By the laws of chance he will probably get some of these correct and will thus achieve a total score higher than that of his colleague. A correction for this effect is often applied to test results. It is based on the following model : suppose a student who takes the test answers correctly on R questions, answers incorrectly on W questions and omits N questions altogether. Suppose further, that all the questions are multiple-choice having m alternative responses only one of which is correct. Then the model says that the student's total of $R+W$ responses are made up of a certain number of cases where the student knows the correct answer and others in which his knowledge is incomplete so that he has to guess. While he guesses fairly blindly, he would be expected to choose the correct response approximately $\frac{1}{m}$ of the time and to choose an incorrect answer $\frac{m-1}{m}$ of the time. The model stipulates that the W wrong responses are all the result of unlucky guesses. Thus, if we take the total number of guesses to be G, we may write the equation

$$W=\frac{m-1}{m} \cdot G$$

or alternatively, the number of guesses,

$$G=\frac{m}{m-1} \cdot W$$

This implies that the number of correct guesses is expected to be $(\frac{1}{m})$th of this, i.e., $\frac{W}{m-1}$ and it is this quantity that is usually subtracted from the total number of correct responses R to give an estimate of the number of items on which the correct response was truly known.

This model is implausible in several respects. Firstly, it does not take into account the case where a pupil acts on the basis of misinformation, and chooses the answer he ' knows ' to be ' right,' even though it is in fact *wrong*. Secondly, it is argued that pupils would usually not be guessing blindly, but rather choosing between two or three alternatives that seem possible to him (which suggests that the formula undercorrects). Thirdly, it is observed in practice that when the formula is applied a considerable number of low ability pupils tend to emerge with negative corrected scores. This suggests that they are not only poor students but also very unlucky ones, against whom the Fates systematically discriminate.

Recent research (Choppin, 1975), however, suggests that if a simple correction of this type is to be applied, then the one offered by the formula is of about the right size. This may well be because the discrepancies from the model noted above tend in practice to cancel each other out.

A Measure of the Tendency to Guess.

Although it is easy to establish that guessing does occur, and that guessing behaviour varies from one student to another, it is not a phenomenon that is easy to measure. This is because when we observe responses being made to items in a test, we have no direct method of telling whether any particular response is the result of a guess or not. To ask the student would be one possible tactic, but it would be difficult to get reliable data on a large scale in this way.

Ziller (1957) has developed an index of the tendency to guess based on the usual model. It is the ratio of the number of guesses estimated to have been made, divided by the estimated total number of items on which the correct answer was not definitely known.

The number of items not definitely known can be estimated by adding the number omitted N, the number of wrong responses W, and the estimated number of correct responses that resulted from guesses

$$\frac{W}{m-1}$$

The estimated number of guesses is the sum of the second and third of these terms, and hence the formula for Ziller's index :

$$G_T = \frac{W + \frac{W}{m-1}}{W + \frac{W}{m-1} + N}$$

and this simplifies to :

$$G_T = \frac{mW}{mW + (m-1)N}$$

This index varies from zero to one, but is undefined in those cases where every item is answered correctly. (i.e., $W=N=O$). For convenience, we multiply the G_T value by 100 and express it as a percentage.

The use of this index will be illustrated by considering some test data recently collected in connection with an international survey of achievement standards.

APPLICATION OF THE METHOD

The Sample.

In 1970, and as part of the British participation in the cross-cultural IEA study of academic standards, a sample of some 160 secondary schools in England and Wales was selected. 14-year-old children in these schools were subjected to a battery of tests covering various aspects of English language and science. The number of children tested was 2,697 and the main results of the study have already been published in various reports (Comber and Keeves, 1973 ; Thorndike, 1973 ; Purves, 1973).

The Tests.

This paper is concerned with the analysis of six separate test instruments that were included in the battery. They were all of the multiple-choice type. Details are given below.

1. *Science Test A* : A 40-item multiple choice achievement test
   Verbal content : Relatively low
   Number of choices per item : 5
   Instructions regarding guessing : You may answer even if you are not quite sure but do not guess blindly.

Sample item :

> Why is it that your body temperature does not fall even though you lose heat continually ?
> A. The blood distributes heat round the body.
> B. Respiration results in the liberation of heat.
> C. Heat is constantly being absorbed from the sun.
> D. Hot meals are eaten regularly.
> E. Warm clothes are good insulators.

2. *Science Test B* : A test of similar content and structure to the one above.
Number of choices per item · 5
Instructions regarding guessing : As for Science A.

Sample item :

> What does an active muscle, that is, a muscle which is doing work, give up to the blood ?
> A. Carbon dioxide.
> B. Oxygen.
> C. Nitrogen.
> D. Vitamin B.
> E. Glucose.

3. *Test of Understanding Science* : A 15-item test designed to assess the degree to which the student understood scientific principles and method.
Verbal content : Comparatively high
Number of choices per item : 5
Instructions regarding guessing : None given

Sample item :

> A scientist predicted that an experiment would come out in a certain way. When he did the experiment, the result was different from what he expected. As a scientist, which of the following would be his most likely reaction ?
> A. "I should not have made a prediction before trying out the experiment."
> B. " I will improve the experiment and make it come out the way I predicted."
> C. " If I had better equipment for the experiment, I would get the right results."
> D. " If I practise long enough, it will come out in the way I want it to."
> E. " Something was wrong either with my prediction, the experiment, or my observations."

4. *Word Knowledge Test* : A 40-item vocabulary measure in which the students had to decide whether the pair of words constituting each item were more nearly synonyms or antonyms.

The test requires little reading

Number of alternatives per item : 2

Instructions regarding guessing : You should attempt every item for which you think you know the answer, but do not guess if you have no idea of the answer.

Sample item :

rapid : sluggish	same opposite

5. *Reading Comprehension Test C* : A test totalling 26 items with high verbal content, designed to assess the ability of pupils to comprehend four separate passages.

Number of alternatives per item : 4

Instructions regarding guessing : You should answer even if you aren't sure. However, do not guess blindly.

Sample item :

> The relation between the first and second paragraphs of this article is that the second paragraph :
> A. Helps explain the results reported in the first.
> B. Gives further detail about the results reported in the first.
> C. Is more specific than the first in the information it provides.
> D. Presents a different point of view from the first.

6. *Reading Comprehension Test D* : A similar test to the one above.

Number of alternatives per item : 4

Instructions regarding guessing : The same as for Reading Comprehension Test C.

Sample item :

> In saying that Ernenek's igloo was ' at the top of the world,' the writer means that it was :
> A. On the icy ocean.
> B. Near the North Pole.
> C. Far from any other home.
> D. Very small and unimportant.

These tests, together with a certain number of other instruments, were administered to the pupils in the sample during two days in March, 1970.

RESULTS

Ziller indices, measuring the tendency to guess, were calculated for each of the six tests already described. The correlation obtained between the indices for the two science tests was 0·64. As these two tests can well be considered parallel, this leads to an estimated reliability for the guessing measure on the combined science tests of 0·78. Similarly a correlation between the two reading comprehension indices of 0·61 leads to a combined reliability estimate of 0·75. In each case these reliability estimates for the measure of guessing are of the same order as those obtained when the tests are considered as measures of achievement in science and reading.

TABLE 1

INTERCORRELATIONS OF THE TENDENCY TO GUESS ON FOUR TYPES OF TEST.

Guessing on :	1	2	3
1. Science (Tests A+B)			
2. TOUS	·35		
3. Word Knowledge Test	·42	·25	
4. Reading Comprehension (Tests C+D)	·33	·24	·33

Table 1 displays the correlations between the two pooled measures, the Test of Understanding Science, and the Word Knowledge test. The entries are all substantial, but not as high as they might be expected to be if the tendency to guess constituted a unidimensional trait. The results appear rather similar to those obtained with measures of school achievement where, although one can usually extract a common factor, one also finds separate abilities in particular subjects.

Guessing indices for each test were also correlated with a verbal ability quotient derived from the Word Knowledge test. The six correlations obtained had a median value of 0·04 which is barely significant, but the fact that all six correlations were both small and positive strongly suggests the generalisation that there does exist a weak correlation between ability and the tendency to guess. High ability students tend to guess slightly more frequently than low ability students.

TABLE 2

SAMPLE SIZE AND THE VERBAL ABILITY QUOTIENTS BY TYPE OF SCHOOL.

	School Type			
	Secondary Modern	Grammar	Comprehensive	Independent
Number of Pupils Tested	1,493	500	418	286
Verbal Ability Quotients	94	111	97	112

When the performance of the pupils is examined according to the type of school attended, the expected differences in verbal ability are found (Table 2). However, it can be seen in Table 3 that the variation in the tendency to guess *does not* mirror the pattern for verbal ability.

TABLE 3

ZILLER INDICES ON SIX SEPARATE TESTS FOR PUPILS AGED 14 IN VARIOUS TYPES OF SCHOOL.

	School Type			
	' Modern '	Grammar	Comprehensive	Independent
Science Test A	89	94	90	88
Science Test B	92	95	93	91
Test of Understanding Science	96	96	96	95
Test of Word Knowledge	80	78	79	78
Reading Comprehension Test C	94	94	93	91
Reading Comprehension Test D	94	93	90	91

A two-way analysis of variance of the data presented in Table 3 yielded highly significant results for both type of school attended and type of test considered. There is also a substantial interaction between the two factors, with the grammar school pupils being particularly likely to guess on the science tests, and the secondary modern pupils on the tests of reading comprehension. In all cases the tendency to guess is high. (The figures shown represent the percentage of occasions on which the students chose to guess rather than omit an item.)

Average results of 60 (France), 62 (West Germany), 79 (Sweden) and 91 (USA) were obtained in an earlier but comparable study for which the English result was 85. It may be noted that guessing is less prevalent among pupils in independent schools, although whether this is due to the more conservative regimes of these schools or the relative lack of exposure of their pupils to multiple-choice testing is still open to conjecture.

Among the tests it is noteworthy that guessing occurs less frequently on the test of word knowledge for which each item has only two alternative responses. In the international study referred to earlier (Choppin, 1975) it was consistently found that the fewer the number of alternatives offered the greater was the tendency to guess. In the present study, the variation in the indices seems to depend less on the number of alternative choices to each question than on the amount of verbal material contained in each question. The more reading a pupil was asked to do, the more likely was he or she to guess.

TABLE 4

ZILLER INDICES BY SEX FOR PUPILS AGE 14.

Test	Boys	Girls	P<*
Science A	91·0	89·3	·01
Science B	93·2	92·1	·05
Understanding Science	95·7	96·2	NS
Word Knowledge	81·6	79·9	·01
Reading Comprehension C	95·2	93·3	·01
Reading Comprehension D	94·1	93·0	·05

* Results of one-tailed t-test.

Analysis of the results for boys and girls separately (Table 4) reveals a significant difference in favour of the boys. The size of the difference is quite small compared with the differences found between types of school, and this may surprise those who consider risk-taking to be associated with the male behavioural stereotype so apparent in our education of young people.

As a similar battery of five tests had been administered to a sample of some 3,600 10-year-old pupils (there was no Test of Understanding Science for this age group), the opportunity was taken to examine the guessing indices by sex for these children also. The results are presented in Table 5.

For this age-group no clear sex differences emerge, but the tendency to guess remains very high.

TABLE 5

ZILLER INDICES BY SEX FOR PUPILS AGE 10.

Tests†	Boys	Girls	P<*
Science Test A	93·6	93·3	NS
Science Test B	94·5	95·4	NS
Test of Word Knowledge	83·6	83·1	NS
Reading Comprehension Test C‡	92·7	94·1	·05
Reading Comprehension Test D	92·7	93·8	NS

* Results of one-tailed t-test.

‡ Difference significant ; P< ·05 difference in favour of girls.

† These tests, though of similar structure to those used with the 14-year-olds, were specifically designed for the 10-year-old sample.

CONCLUSION

This study further establishes the utility of the index first proposed by Ziller (1957) to measure an individual's tendency to guess on a particular instrument.

Correlations between scores suggest that individuals behave with a certain degree of consistency when faced with different tests. The value of the index, however, does not remain constant as the test changes. In general, increasing the verbal content of a question would seem to raise the level of the index. Reducing the number of alternative answers to each question also has this effect. Guessing is more common in England than in other West European countries, and much more prevalent than in Eastern Europe where it is actively discouraged.

Rather surprisingly, the index appears to be almost uncorrelated with ability. Guessing is almost always a better strategy than is omitting items—even in cases where a penalty for guessing in the form of the usual ' correction ' applies. There is, however, little evidence that high ability pupils use this strategy more frequently than those less able. At age 14, boys guess slightly more frequently than do girls, but this difference was not found in a similar study of 10-year-olds.

The Ziller index, which can be derived in all cases where the usual correction for guessing is applied, may, in future, prove a useful tool in investigations of test-taking behaviour.

REFERENCES

CHOPPIN, B. (1975). *The Correction for Guessing on Objective Tests.* Stockholm : IEA (in press).

COMBER, L. C., and KEEVES, J. P. (1973). *Science Education in Nineteen Countries : International Studies in Evaluation I.* New York : Wiley.

DIAMOND, J., and EVANS, W. (1973). The correction for guessing. *Rev. educ. Res.*, 43, 181-191.

PURVES, A. C. (1973). *Literature Education in Ten Countries : International Studies in Evaluation III.* New York : Wiley.

THORNDIKE, R. L. (1973). *Reading Comprehension in Fifteen Countries : International Studies in Evaluation II.* New York : Wiley.

ZILLER, R. C. (1957). A measure of the gambling response set in objective tests. *Psychometrika* 22, 289-292.

(*Manuscript received* 10*th October*, 1974)

Br. J. educ. Psychol., 45, 214-226, 1975

THE FACTORIAL STRUCTURE OF WECHSLER'S PRE-SCHOOL AND PRIMARY SCALE OF INTELLIGENCE AT SUCCESSIVE AGE LEVELS

BY B. N. MUKHERJEE
(*Council for Social Development, New Delhi*)

SUMMARY. A factor analysis of the correlation matrix for the subtests of WPPSI was carried out for each of the age groups between 4 and 6½ years. The data were obtained from the WPPSI standardisation samples. In each analysis, three factors were extracted and two of these, after equimax rotation, were found to give patterns of loadings which were similar for all age groups. The rotated factors were identified as verbal intelligence, space performance and comprehension. Order analysis of the same data tended to support the notion that these factors could be also arranged into a circular order along with various tests of WPPSI. Using a likelihood ratio test for the Guttman ' circumplex ' structure, the correlation matrix for the 4-year age group was analysed and the result provides clear support for the postulated circular arrangement of the tests. The circumplex ordering of WPPSI suggests a systematic ordering according to the law of neighbouring variables and is found to be more parsimonious and less subjective in interpreting than the results emerging from the traditional factor analysis.

INTRODUCTION

THE present study was originally undertaken in order to examine the possibility of age-related trends in the pattern of mental abilities as reflected in the factorial structure of the Wechsler Preschool and Primary Scales (WPPSI), which had been administered to children aged 4 and 6½ years. For this purpose, the WPPSI data were factored by the iterative principal-axes method and three factors were retained for rotation by the equimax method.

The study reported here aimed at answering the following questions :

(*a*) With the principal-axes method can the first of the factors so obtained be identified with general intelligence and, if so, is this factor the same for all the age groups ?

(*b*) Does this general factor account for a decreasing amount of common variance as the age level increases ?

(*c*) With the method of equimax rotation how far do the factors obtained at one age level correspond with factors obtained at another ? If they are the same, do the factor loadings show any monotonic trend as a result of age change ? The final, and most important question was :

(*d*) Do the WPPSI subtests conform to a circular ordering or ' circumplex ' (Guttman, 1954) ?

The circumplex theory of Guttman (1954) postulates that qualitatively different traits in a given domain can have an order among themselves without a beginning or an end. "A set of variables obeying such a law will be called a *circumplex* to designate a ' circular order of complexity ' " (Guttman, 1954, p. 260). In an intercorrelation matrix which has a circumplex structure, " the largest correlations are along the main diagonal, and increase again towards the upper and lower right-hand corners. More importantly, the correlations located along each diagonal parallel to the main diagonal are about equal " (Mukherjee, 1973). In short, the intercorrelation matrix of circumplex structure shows the contiguity pattern of a symmetric ' circulant ' matrix.

Although Guttman (1954) did not attempt to limit the methods by which a circumplex order can be discovered, he explicitly advocated the method of visual inspection (Guttman, 1957). Recently, Guttman (1966) has developed the method of 'order analysis.' In the present study, factor analysis, order analysis and a new multivariate technique known as covariance structural analysis (Mukherjee, 1970b, 1973) have been used together to examine whether or not the WPPSI subtests generate a 'circumplex' structure in the intercorrelation matrix.

METHODS

Subjects.

The data analysed were obtained from children of between 4 and 6½ years old in the standardisation samples reported in the WPPSI Manual (Wechsler, 1967). Each age group (sample) consisted of 100 boys and 100 girls, including rural and urban children roughly in proportion to the number of such children in the US population. In addition, each sample is claimed to be representative in terms of 1960 US Census data with reference to geographical region, rural-urban residence, and father's occupation. Only white children were covered in the standardisation study.

The WPPSI Scale consists of 11 subtests and eight of these subtests are downward extensions and adaptations of the subtests of the Wechsler Intelligence Scale for Children (WISC). As in the WISC, the raw scores on each WPPSI subtest are transformed into standard scores with a mean of 10 and a standard deviation of 3 within each quarter-year group. Thus, all subtest scores are expressed in comparable units. Because of the uniform scale of measurement at each age group for all the 11 subtests, the results of factor analysis will lead to essentially the same conclusions whether the correlation matrices are analysed or the variance covariance matrices. Since the standard deviation of each subtest remains more or less constant with increasing age, any change in the factorial structure from one age to another can, therefore, be attributed to age differences rather than to changes in the unit of measurement.

The six intercorrelation matrices (each of 11×11 order) as given in the WPPSI *Manual* (Wechsler, 1967) were analysed separately by means of the principal-factor method (Harman, 1967) with iteration. The computations were carried out using the PA-2 factor analysis programme of the *Statistical Package for the Social Sciences* (SPSS) written by Nie, Bent and Hull (1970). This program first calculates the squared multiple correlations (SMC) between each variable and the remaining (p-1) variables of the non-singular intercorrelation matrix and uses these SMC values as the communality estimates, which are then improved iteratively until the differences in the two sets of communality estimates become negligible.

An examination of the latent roots showed a large difference between the first and second factors. This fact, together with Kaiser's criterion, suggested that it might be desirable to retain only the first two factors. However, the following criteria suggested that three principal factors would be needed to account for the non-random common variance :

(1) The residuals after extraction of three factors were found to be so small that it was decided to be unwise to extract further factors ; and (2) the scree test (Cattell, 1966a, 1966b) also suggested that after the third principal factor, the extraction of the remaining factors was not necessary. The three factors accounted for as much as about 51 to 57 per cent of the total variance in each analysis.

The three principal factors obtained for each intercorrelation matrix were then rotated to an orthogonal simple structure by using the equimax method. This method of rotation can be thought of as a compromise solution of Kaiser's (1958) varimax and Neuhaus and Wrigley's (1954) quartimax criteria. Instead of concentrating either on simplification of the rows or on simplification of the columns, it attempts to accomplish some of each by making a maximum number of entries (values) in each row or column of the factor matrix as close as possible to zero.

RESULTS

When percentages of the total variance attributable to successive principal factors were calculated from the latent roots, these percentages were found to be very similar for all age groups. This finding supports the notion of the structural invariance of the WPPSI intercorrelation matrices at different age levels.

In each age group, the principal-axes solution obtained for the various intercorrelation matrices yields a first general factor which has higher saturation than any other factor in every test, which can thus be interpreted as a *general intelligence* factor. Inspection of the principal factor loading matrix (omitted through lack of space) reveals that the pattern for this general factor is essentially the same in every age group. In order to decide whether or not the influence of this general factor shows a consistent diminution with increasing age, the percentage of common-factor variance explained by the general factor at each age level was examined. It was found that between 4½ years and 6 years there is a discernible increasing trend in the percentage of common variance explained by the first general factor. At age 4 years, this percentage is high and after falling at 4½ years it gradually rises until its final decrease at 6½ years. Since the difference in these proportions from one age level to another is not substantial, the evidence of this research in itself is not entirely conclusive.

Interpretation of Rotated Factors.

The factor structure after equimax rotation is shown in Table 1. The first rotated factor, in each analysis, can be interpreted in terms of *verbal intelligence* since all subscales of the verbal part of the WPPSI have high loadings on it. The only exceptions to this finding are the Arithmetic test at ages 4 and 6½ and the Sentence test at age 6 years.

Interpretation of the second rotated factor appears to be quite straightforward. Except for ages 4½ and 5 years, the tests which exhibited high loadings of second rotated factor are supposed to measure spatial organisation. Since the second rotated factor is found to have negligible loadings on each of the first five verbal tests listed on the top of the factor matrix (see Table 1) but has consistently high loadings, for each age group excepting the 4½ and 5-year groups on most of the performance tests, the factor can be labelled *space performance* following Maxwell (1959). Alternatively, as most of the tests showing high projections on the 'space performance' factor require visual rather than verbal processes in solution of the problems, this factor can be described as 'visual ability.'

At the 4½ and 5-year age levels, the second rotated factor is tentatively interpreted as '*comprehension*,' since this factor is found to have high loadings on Information, Vocabulary, Arithmetic and Picture Completion at these two age levels. The third rotated factor at these age levels can be identified as the space performance factor, since all the performance subscales have high loadings on this factor.

TABLE 1

EQUIMAX FACTOR LOADINGS FOR SUCCESSIVE AGE GROUPS FROM 4 TO 6½ YEARS.

Test Name	Rot. Factors at 4 years			Rot. Factors at 4½ years			Rot. Factors at 5 years			Rot. Factors at 5½ years			Rot. Factors at 6 years			Rot. Factors at 6½ years		
	V	S	C	V	S	C	V	S	C	V	S	C	V	S	C	V	S	C
Information	68	28	33	51	60	22	62	32	30	68	35	35	63	30	34	64	23	41
Vocabulary........	55	28	34	47	59	10	60	26	26	60	16	47	67	28	30	61	33	23
Arithmetic	24	40	56	49	32	40	53	53	21	54	42	41	42	40	40	30	37	77
Similarities	55	10	54	62	34	23	54	9	19	61	12	26	50	25	39	66	18	31
Comprehension ..	63	25	37	70	36	15	68	17	35	66	26	32	64	20	40	72	20	22
Sentences	41	19	61	69	22	28	52	29	17	64	36	9	26	24	78	50	8	50
Animal House	41	27	24	33	26	40	18	42	46	28	50	29	13	58	34	17	35	30
Picture Completion	42	49	22	15	70	35	25	20	68	28	37	47	41	52	22	28	45	24
Maze	14	74	17	5	10	64	21	36	52	13	38	68	36	54	31	11	71	9
Geom. Design .	26	61	18	25	17	57	19	63	39	11	76	32	21	72	24	13	53	33
Block Design	14	47	40	19	28	53	16	76	27	29	63	33	43	53	24	28	54	32
Contribution of Factor	2·15	1·88	1·67	2·30	1·79	1·66	2·24	1·90	1·55	2·60	2·05	1·66	2·30	2·19	1·68	2·30	1·77	1·58
Percentage of Total Variance	20·0	17·0	15·0	21·0	16·0	15·0	20·0	17·0	14·0	24·0	19·0	15·0	21·0	20·0	15·0	21·0	16·0	14·0

Decimal points from all loadings have been omitted.

Factor V is identified as *Verbal Intelligence*, Factor S is interpreted as *Space Performance*, and Factor C is tentatively labelled *Comprehension* for all age groups excepting the 6½-year where the factor emerges more as a numerical ability.

For the other age groups, the interpretation of the third rotated factor appears to be somewhat difficult. The comprehension factor is ambiguous, not only because of low loadings, but also because of the tests on which it is demonstrated. At ages 4 and 6 years, the third rotated factor has moderately high loading on the Arithmetic and Sentence tests, both of which call for comprehension. For the remaining groups, namely, 5½ and 6½ years, the third rotated factor can be tentatively identified as '*comprehension*' since tests such as Vocabulary, Arithmetic, Comprehension, Sentences, Picture Completion and Maze appear high on this factor.

Change in the Factorial Structure.

Having interpreted the rotated factors in the above manner, the amount of similarity between the factorial structure of two or more factor matrices can be assessed informally without using any rigorous statistical tests. As has been mentioned earlier the two rotated factors, those identified with *verbal intelligence* and *space performance*, are defined almost by the same tests in both higher and lower age levels. This indicates that the pattern shown by these factors is much the same in all the age groups. Although the space performance factor emerges as a third, rather than a second, rotated factor at the 4½ and 5-year level, the variables defining it still approximate closely to the group of performance tests which defined this factor at other age levels. The change of sequence at the above two age levels only meant a small reduction (about 2 to 3 per cent) in the percentage of common variance contributed by the 'space' factor at the 4½ and 5-year age levels.

In the case of the rotated factor identified as 'comprehension' there is some change from one age level to another. Firstly, the factor at ages 4 and 6½ showed such high loadings on the Arithmetic test that it can also be interpreted as *numerical ability*. At 4½ years the Picture Completion test shows the largest projection on this factor, followed by the tests of Information and Vocabulary. The comprehension factor at age 5½ years is defined by Maze, Vocabulary, Arithmetic and Picture Completion. At 6½ years Sentences, Comprehension, Arithmetic and Similarities show the highest loadings, while Picture Completion is less prominent.

At the 6½-year level, the third rotated factor clearly emerges as a 'numerical' factor although tests expected to be tapping 'comprehension,' such as Sentences, Information, and so on, also exhibit moderately high loadings on this factor. The emergence at the 6½-year level of a specialised factor which seems to be associated with numerical ability suggests that the gradual differentiation of ability as hypothesised by Wright (1939), Garrett (1946), Anastasi (1948), Burt (1954) and Mukherjee (1962) might start as early as the beginning of primary school level. Further studies of the longitudinal type are needed to substantiate the above hypothesis.

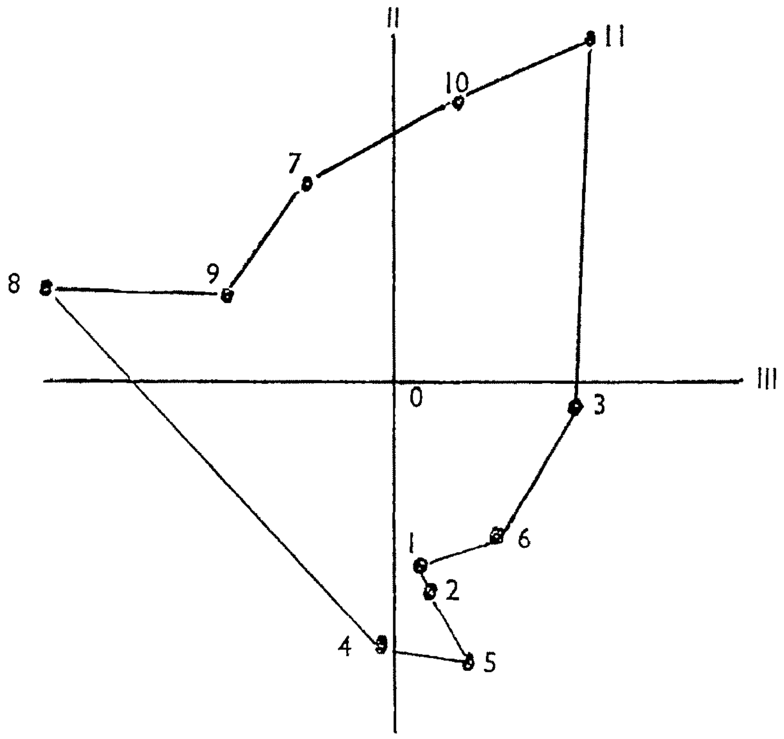
An Alternative Interpretation.

Instead of indulging in the subjectivity (Mukherjee, 1973) of interpreting factors in this way, it is better to examine the intercorrelation matrices for evidence of any discernible pattern. If this is found, then we need to determine whether the tests can be ordered in terms of hypothetical constructs. In the case of WPPSI intercorrelation matrices, an attempt was made to see if the various tests could be mapped near the circumference of a unit circle. To this end, the loadings of the second principal factor (ordinate) on each of the 11 test variables were plotted against their respective loading on the third principal factor (abscissa). The plot on the rectangular co-ordinate system shows the relative position of each variable in pairs of axes.

As an illustration, the relative position of the 11 tests of WPPSI administered at age 5 years are shown in Figure 1 on a two dimensional space represented by second and third principal factors. In examining the graph the following three aspects should be noted : (*a*) the relative distance of a variable from the two axes ; (*b*) the direction of a variable in relation to the axes (it may indicate either a positive or a negative factor loading), and finally (*c*) the clustering of variables and their relative position to each other.

FIGURE 1

PLOTTING OF FACTOR LOADINGS OF FACTOR II (Ordinate) AGAINST FACTOR LOADINGS OF FACTOR III (ABSCISSA) FOR ALL THE ELEVEN SUBSCALES OF WPPSI (5-YEAR-OLD AGE GROUP).



If each test in Figure 1 is joined to the adjoining ones, it is seen that none is close to the origin and all the tests are near the periphery of an ellipse. The finding suggests the possibility of ordering the WPPSI subscales on the circumference of a circle. In this age group the same ordering emerges if the loadings of rotated Factor 1 are plotted against rotated Factor 3.

The Circular Order.

Suppose a set of p variables are such that they can be ordered into a circular sequence. Then, according to the 'law of neighbouring' any specified variable, say X_1, mapped on the circle will have its largest correlations with the two neighbouring variables, viz. with X_2 and X_p. Its correlations with remotely located neighbouring variables on the clockwise and anti-clockwise direction will fall off monotonically as a function of their distance on the circumference of the circle.

TABLE 2

PRODUCT MOMENT INTERCORRELATIONS* AMONG SUBTESTS OF WPPSI FOR 5-YEAR AGE LEVEL
(From Wechsler, 1967, N=100 Boys and 100 Girls.)

Test Name and Serial No.		Test Name and Serial Number (Original)										
		Voc. 2	Inf. 1	Sen. 6	Arith. 3	B.D. 11	Geom 10	An.H. 7	Maz. 9	Pic. 8	Comp 5	Sim. 4
Vocabulary	2	—										
Information	1	·56	—									
Sentences	6	·41	·43	—								
Arithmetic	3	*·55*	*·57*	·48	—							
Block Design	11	·35	·42	·35	·54	—						
Geom. Design	10	·38	·42	·32	·50	·63	—	Symmetric				
Animal House	7	·33	·38	·35	·40	·46	·49	—				
Mazes	9	·33	·42	·28	·43	·45	·46	·42	—			
Pict. Comp.	8	·40	·41	·30	·39	·38	·43	·45	·49	—		
Comprehension	5	·56	*·59*	·43	·49	·35	·40	·35	·39	·43	—	
Similarities	4	*·34*	*·40*	·41	·34	·22	·26	·19	·24	·29	·49	—
Mean...............		9·9	9·9	9·9	9·8	9·9	9·9	9·9	10·0	10·0	9·9	9·9
Standard Deviation ..		3·0	3·1	3·1	3·1	3·1	3·1	3·0	3·1	2·9	3·1	3·1

*Coefficients which are in italic are either larger or smaller than the values expected under the circumplex hypothesis.

Table 2 presents the intercorrelations for the 5-year age group among the 11 subscales of WPPSI after they have been ordered in the sequence as shown in Figure 1. Examination of the table reveals that the circular order requirements are satisfied to a great extent. While the correlations along the main diagonal tend to be the largest, most of the correlation coefficients between Test 4 (Similarities) and other variables shown in the last row of Table 2 are lower than their expected values. Such coefficients appear in Table 2 in italics. However, with the exception of a few coefficients, the values which run parallel to the main diagonal are found to be almost equal. If we drop from the matrix, Test 1 (Information), Test 4 (Similarities) and Test 7 (Animal House), then many of the reversals which are in italic in Table 2 will be removed. With the deletion of the above variables, the circumplex structure in the case of the 4-year age group emerges clearly, as is seen from Table 3. Furthermore, with these three variables excluded, the WPPSI intercorrelation matrix for most of the other age groups also shows the circumplex structure. The eight tests of WPPSI that fit into a circular order are presented both in Table 3 (for 4-year age group) and in Figure 2.

TABLE 3

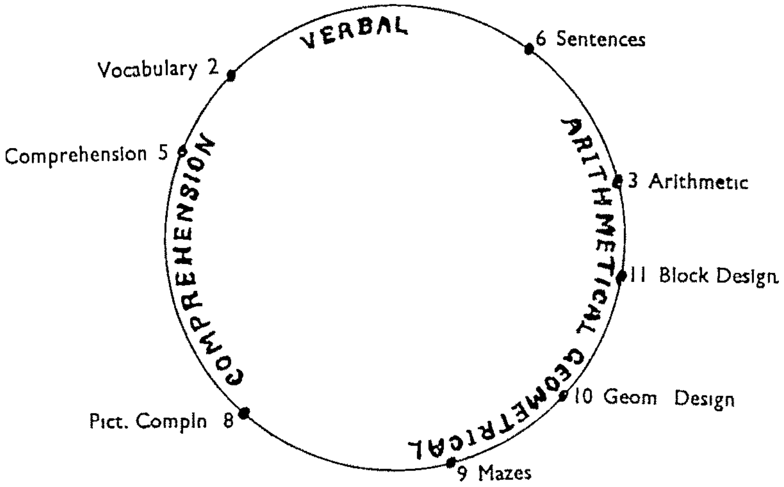
PRODUCT MOMENT INTERCORRELATIONS AMONG EIGHT TESTS OF WPPSI WHICH FIT INTO A CIRCULAR ORDER.
(From Wechsler, 1967, N=100 Boys and 100 Girls, 4-Year Age Level.)

Original Serial No.	2	6	3	11	10	9	8	5
Name of the Test	Vocab.	Sent.	Arith.	Bl. D	Geo.D	Mazes	Pic. C.	Comp.
Vocabulary	1·00	·51	·41	·38	·38	·31	·45	·57
Sentences................	·51	1·00	·52	·36	·34	·33	·35	·53
Arithmetic	·41	·52	1·00	·46	·44	·39	·42	·48
Block Design	·38	·36	·46	1·00	·35	·43	·42	·32
Geom. Design	·38	·34	·44	·35	1·00	·53	·40	·38
Mazes	·31	·33	·39	·43	·53	1·00	·47	·34
Pict. Comp.	·45	·35	·42	·42	·40	·47	1·00	·47
Comprehension	·57	·53	·48	·32	·38	·34	·47	1·00
Column Sum	4·01	3·94	4·12	3·72	3·82	3·80	3·98	4·09
Mean	9·9	9·9	9·8	9·9	9·9	10·0	10·0	9·9
Standard Deviation	3·0	3·1	3·1	3·1	3·1	3·1	2·9	3·1

It may be noted that with few exceptions each row of Table 3 has approximately the same entries as the preceding row, except that the entries are moved one space to the right with the end entry in one row moving to the beginning of the next row. Furthermore, the row or column totals are not very different from each other. This indicates that the eight subtests of WPPSI have an approximately equally-spaced circumplex pattern (Guttman, 1954).

FIGURE 2

A CIRCUMPLEX ORGANISATION OF MENTAL ABILITIES BASED ON WPPSI DATA.



An idea as to the nature of the circumplex involved in the above example is given by the inverse of the intercorrelation matrix of Table 3. The multiple regression weights (standardised) and the multiple correlation for predicting each subtest from the remaining seven subtests are displayed row-wise in Table 4, further to verify the law of circular ordering.

TABLE 4

STANDARDISED MULTIPLE REGRESSION WEIGHTS AND MULTIPLE CORRELATION COEFFICIENTS FOR PREDICTING EACH SUBTEST OF WPPSI FROM THE REMAINING SEVEN SUBTESTS (4-Year Age Level).

Criterion Variable	Weights of the Predictor Variables								
	Vocab.	Sent.	Arith.	Bl. D.	Geo.D.	Mazes	Pic. C.	Comp	R
Vocabulary	—	0·2367	–0·0097	0·1410	0·1210	–0·0555	0·1675	0·3002	·659
Sentences	0·2352	—	0·2664	0·0639	0·0026	0·0633	–0·0357	0·2038	·645
Arithmetic	–0·0323	0·2559	—	0·2481	0·1796	0·0316	0·0772	0·2227	·661
Block Design	0·1238	0·0605	0·2075	—	0·0180	0·1876	0·1453	–0·0645	·576
Geom. Design	0·1130	0·0042	0·1676	0·0174	—	0·3495	0·0606	0·0436	·611
Mazes	–0·0514	0·0615	0·0313	0·2079	0·3633	—	0·2348	0·0167	·631
Pict. Comp.	0·1568	–0·0363	0·0854	0·1552	0·0599	0·2308	—	0·1718	·623
Comprehension ..	0·3220	0·2474	0·1635	–0·0636	0·0660	0·0216	0·2070	—	·684

The pattern of the multiple regression weights clearly indicates that the neighbouring variables which are adjacent to each other in Table 3 contribute significantly in the prediction of each other. The regression weights, therefore, show more or less the same pattern as the intercorrelation matrix. Moreover, the multiple correlations are approximately of the same magnitude, again suggesting that the eight subtests of WPPSI are not very unequally spaced.

In view of the above results of order analysis, it can be then said that the WPPSI subscales tend to exhibit more or less an equally spaced circular order and this interesting finding has never emerged in the previous factor analytic studies of WPPSI (Silverstein, 1969). The WPPSI subtests which form the hypothesised sequential order or the circumplex (see Figures 1 and 2) are as follows : Vocabulary, Sentences, Arithmetic, Block Design, Geometric Design, Mazes, Picture Completion and Comprehension. It seems possible then that the various tests of WPPSI can be mapped near the circumference of a circle running from verbal ability, through numerical, geometrical, visual and comprehension and back to verbal again, as has been shown in Figure 2. The name of the primary abilities appear in capital letters inside the circle whereas the idealised location of the eight tests of WPPSI are shown with their corresponding serial number on the circumference of the circle.

Further Evidence for the Circular Order.

Direct statistical evidence supporting the circular order hypothesis regarding the eight WPPSI subtests listed in Table 3 was secured by employing a likelihood ratio test (see Mukherjee, 1970a, 1970b, 1971) derived for the Guttman ' circumplex ' structure. The likelihood ratio test was used for testing the null hypothesis that in the population of 4-year-old American (white) children, the eight subtests of the WPSSI having a known common variance would show equality of correlations located along each diagonal parallel to the main diagonal—in other words that they would fulfil the conditions for a circumplex structure as outlined earlier. From a computation of the ' lambda criterion ' (see Appendix) we are obliged to accept the null hypothesis and conclude that the eight subtests of WPSSI at the 4-year age level do form a circumplex structure when they are arranged in the same order as shown in Table 3. Had the sample size been smaller (N<200), we would be even more confident of the above inference. The intercorrelation matrices of the eight WPPSI subtests for other age groups have been subjected to the same type of covariance structural analysis as illustrated above and in each case the results appear to justify the acceptance of the circumplex structure hypothesis.

DISCUSSION

The present factor-analytic investigation of WPPSI does not show any substantial change as a function of age in the nature and composition of the first principal factor which was interpreted as general intelligence. In each analysis, three factors were extracted and two of these, after rotation, were found to give patterns of loadings which were similar for all the six age groups studied. These two rotated factors can be identified as verbal intelligence and space performance. The study generally supports the notion that the so-called performance tests of the WPPSI form a cluster which can be interpreted as a 'performance' factor since in each analysis the rotated factor labelled 'space performance' showed high loadings on all performance tests of WPPSI. In most cases, the 'space performance' factor is found to have negligible loading, on most of the first five subtests which form the verbal part of WPPSI. Thus as far as the WPPSI is concerned, the study supports Wechsler's contention that "the tests identified as verbal and performance differ as these labels indicate" (Wechsler, 1949, p. 5). However, it seems questionable to think of them as defining two relatively independent factors of intelligence, since the composite verbal and performance scores correlate very highly with each other according to the WPPSI Manual, the product-moment correlations ranging from 0·62 to 0·71.

The present study also identified a third common factor which, after rotation, was interpreted as a comprehension ability for all age groups except one. While there were substantial differences in the composition of this comprehension factor at different age levels, the factor emerges clearly as a numerical factor at the 6½-year level.

In contrast to the findings discussed above, order-analysis of the WPPSI data followed by the sophisticated technique of covariance structural analysis clearly indicated the possibilities of ordering the various tests of the WPPSI at each age level near the circumference of a circle running from verbal ability through numerical and geometrical abilities (as well as comprehension ability) and back to verbal ability again. This technique circumvents the well-known problems associated with factor rotation by making use only of the principal factor loadings. Besides avoiding the subjective interpretation of factors, order analysis also presents a more parsimonious description of intellectual structure. All the factors emerging from the factor analysis of WPPSI data can be plotted simultaneously with all the eight tests at different points on the periphery of a circle as shown in Figure 2. Thus, in comparison to the results of factor analysis, the findings emerging from order analysis display more clearly the properties of descriptive efficiency and statistical invariance which according to Peterson (1965) offer maximal promise for predictive utility.

Although Guttman (1954) did not limit the dimensionality of his circumplex model, the data considered here are well ordered within a two-dimensional space. A sequential order of variables can be obtained by ordering a correlation matrix according to the 'law of neighbouring' and a spatial representation of this sequential order can be generated by factor analysis. For the purpose of understanding the domain of intelligence the sequential ordering of abilities revealed by plotting the factor loadings seems more interesting than merely extracting and identifying the common factors, as is done in traditional factor-analytic studies. The plot of factor loadings could be interpreted either in terms of the ordered set of WPPSI variables or in terms of common factors identified as space performance and comprehension. The former interpretation is conceptually more useful since knowledge of neighbours, polar opposites and unrelated variables provides a nomological network which contributes to the construct validity of different tests constituting the whole battery (Cronbach and Meehl, 1955).

The idealised mapping of WPPSI variables into a circumplex shows that verbal ability is located farthest from the space performance or visual factor. As such, it should have the lowest correlation with the spatial factoı. The correlation with arithmetic ability increases as tests become more verbal or more spatial or visual in nature. Arithmetic ability should correlate least, but not negatively, with what is opposite it on the circle—something between the verbal and the space factor. In the case of WPPSI, this ' something ' appears to be a comprehension factor. Thus, the present order analysis of the WPPSI tests succeeds in portraying in two dimensions the circumplex structure of the intercorrelations among 11 tests of WPPSI despite the fact that the conventional factor analysis of the same data for the 5-year age group requires three dimensions.

The ordering of primary mental abilities based on the WPPSI data as found in this study is very similar to the ordering observed by Guttman (1954, p. 336 ; 1957, p. 401) in the intercorrelations reported by Thurstone and Thurstone (1941) for selected tests of intelligence administered to 710 Chicago school children. A subsequent study will examine how far the same circular ordering is found in the WISC and Wechsler's Adult Intelligence Scale intercorrelation matrices.

Apart from answering the question of how the structure of mental abilities differs with increasing age and how the structural changes affect the circular ordering of the Wechsler tests, such studies, including the present one, have direct relevance to education. The theory of circumplex organisation of mental abilities as substantiated here is not as narrow as the conception of ' general ' intelligence. It might thus be expected to encourage among teachers the possibility of launching appropriate programmes for the development of various kinds of abilities right from the preschool stage. The mental abilities identified by the WPPSI should stimulate educational planners to think of some such blending of curricular measurement material as suggested by the circular notion of abilities.

Factors measured by WPPSI such as verbal understanding, quantitative reasoning and visual (figural analytic) thinking are basic academic skills relating comprehension power to three functionally distinct media in the academic curriculum. For school purposes they appear highly predictive and more meaningful than a dichotomous scheme that classifies intelligence as verbal and performance. The number of differential abilities as measured by WPPSI and as identified in Figure 2 is small. The abilities are neither so general as measured by the Stanford Binet test of intelligence nor so specific as those measured by Guilford's tests. Furthermore, each of the abilities identified in the study can be seen as reflections of the environmental field which varies as educational and cultural pressures dictate.

From the theoretical point of view, the proposed circular ordering of mental abilities serves to link together and organise these abilities into a unified system—a clearer nomological network than the factor-theoretic conception of intelligence. Instead of dealing with idiosyncratic details, as in Guilford's ' Structure of Intellect ' model, the circumplex theory directs attention to the sequential order of test variables and does not divide the matrix of intercorrelations among these variables into discrete clusters or factors. For the researchers, the problem of completing the existing gaps in the circle of mental abilities may be an important challenge.

REFERENCES

Anastasi, A. (1948). The nature of psychological ' traits.' *Psychol. Rev.*, 45, 127-138.

Burt, C. (1954). The differentiation of intellectual ability. *Br. J. educ. Psychol.*, 24, 76-90.

Cattell, R. B. (1966a). The meaning and strategic use of factor analysis. In Cattell, R. B. (Ed.), *Handbook of Multivariate Experimental Psychology*. Chicago : Rand McNally.

Cattell, R. B. (1966b). The scree test for the number of factors. *Multiv. Behav. Res.*, 2, 245-276.

Cronbach, L. J., and Meehl, P. E. (1955). Construct validity in psychological tests. *Psychol. Bull.* 52, 281-302.

Garrett, H. E. (1946). A developmental theory of intelligence. *Am. Psychol.*, 1, 372-378.

Guttman, L. (1954). A new approach to factor analysis : the radex. In Lazarsfeld, P. (Ed.), *Mathematical Thinking in Social Sciences*. Glencoe, Ill. : Free Press, 258-348.

Guttman, L. (1957). Empirical verification of the radex structure of mental abilities and personality traits. *Educ. psychol. Measur.*, 17, 391-407.

Guttman, L. (1966). Order analysis of correlation matrices. In Cattell, R. B. (Ed.), *Handbook of Multivariate Experimental Psychology*, 438-458. Chicago : Rand McNally.

Harman, H. H. (1967). *Modern Factor Analysis*. Chicago : University of Chicago Press.

Kaiser, H. F. (1958). The varimax criterion for analytical rotation in factor analysis. *Psychometrika*, 23, 187-200.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educ. psychol. Measur.*, 20, 141-151.

Maxwell, A. E. (1959). A factor analysis of the Wechsler Intelligence Scale for Children. *Br. J. educ. Psychol.*, 29, 237-241.

Mukherjee, B. N. (1962). The factorial structure of aptitude tests at successive grade levels. *Br. J. statist. Psychol.*, 15, 59-69.

Mukherjee, B. N. (1970a). A likelihood ratio test for the Guttman quasi-circumplex covariance structure. Paper presented at a seminar, Department of Measurement and Evaluation, Ontario Institute for Studies in Education, Ontario, Canada, on April 3rd, 1970.

Mukherjee, B. N. (1970b). Likelihood ratio tests for statistical hypotheses associated with patterned covariance matrices in psychology. *Br. J. math. statist. Psychol.*, 23, 98-120.

Mukherjee, B. N. (1971). A method of fitting the quasi-circumplex symmetrically. Paper presented at the Spring meeting of the Psychometric Society, St. Louis, Missouri, USA, on April 25th, 1971.

Mukherjee, B. N. (1973). Analysis of covariance structures and exploratory factor analysis. *Br. J. math. statist. Psychol.*, 26, 425-455.

Neuhaus, J. O., and Wrigley, C. (1954). The quartimax method : an analytical approach to the orthogonal simple structure. *Br. J. math. statist. Psychol.*, 7, 81-91.

Nie, H. H., Bent, D. H., and Hull, C. H. (1970). *Statistical Package for the Social Sciences*. New York : McGraw-Hill.

Peterson, D. P. (1965). Scope and generality of verbally defined personality factors. *Psychol. Rev.*, 72, 48-59.

Silverstein, A. B. (1969). An alternative factor analytic solution for Wechsler's Intelligence Scales. *Educ. psychol. Measur.*, 29, 763-767.

Thurstone, L. L., and Thurstone, T. G. (1941). Factorial studies in intelligence. *Psychometric Monograph, No.* 2. Chicago : University of Chicago Press.

Wechsler, D. (1949). *Wechsler Intelligence Scale for Children*. New York : Psychological Corporation.

Wechsler, D. (1967). *Manual for the Wechsler Preschool and Primary Scale of Intelligence*. New York : Psychological Corporation.

Wright, R. E. (1939). A factor analysis of the original Stanford-Binet Scale. *Psychometrika*, 4, 209-220.

(*Manuscript received* 17*th July*, 1974)

APPENDIX

The Likelihood-Ratio Test.

Details of the computational steps involved in obtaining the maximum likelihood estimates are described elsewhere (Mukherjee, 1970a, 1970b, 1971) along with the derivation of the likelihood ratio test for the circumplex structure. In order to construct the likelihood ratio test of the null hypothesis that $\underset{\sim}{\Sigma}$ can be written as a symmetric circulant matrix against the alternative hypothesis that $\underset{\sim}{\Sigma}$ is a positive definite matrix, one computes the lambda criterion, λ, given by

$$\lambda = \left[\frac{|\underset{\sim}{\hat{\Sigma}_o}|}{|\underset{\sim}{S}|}\right]^{-N/2},$$

where N is the sample size, $\underset{\sim}{\hat{\Sigma}_o}$ is the maximum likelihood estimate of $\underset{\sim}{\Sigma}$ under the stated null hypothesis. The terms $|\underset{\sim}{S}|$ and $|\underset{\sim}{\hat{\Sigma}_o}|$ refer to the determinants of the sample variance covariance matrix and the estimated variance covariance matrix, respectively. The statistic,

$$-2 \log_e \lambda = N \log_e \frac{|\underset{\sim}{\Sigma_o}|}{|\underset{\sim}{S}|}$$

is asymptotically distributed as χ^2, where the degrees of freedom (d.f.) are given by

$$\begin{aligned} df &= \frac{(p^2-p)}{2} - g \\ &= \frac{(8^2-8)}{2} - 4 = 24 \text{ (for the given sample)} \end{aligned}$$

where 'g' represents the number of estimable elements of the p×p variance covariance matrix, Σ_o. Therefore, the results of the likelihood-ratio test can be written as :

$$\begin{aligned} \chi^2_{24} &= N \log_e \frac{|\underset{\sim}{\hat{\Sigma}_o}|}{|\underset{\sim}{S}} \\ &= 200 \log_e \frac{(0{\cdot}0727921)}{(0{\cdot}0615128)} \\ &= 200 \log_e 1{\cdot}18 \\ &= 200\,({\cdot}1655) = 33{\cdot}10 \end{aligned}$$

which for 24 df is not significant even at the ·10 level.

RESEARCH NOTES

Br. J. educ. Psychol., 45, 227-231, 1975

ATTITUDINAL, SITUATIONAL AND MENTAL HEALTH CORRELATES OF ACADEMIC ACHIEVEMENT AT THE UNDERGRADUATE UNIVERSITY LEVEL

By K. M. BANRETI-FUCHS

*(Laurentian University, Sudbury, Ontario, Canada)**

Summary. The relationship between levels of academic achievement of first-year university students and various attitudinal, situational and mental health factors was examined. Using the regression equation method a sample of 222 first-year students were assigned to a high-achieving, average-achieving, or low-achieving group. The classification of each student was based on the relationship between his actual grade point average (GPA) in introductory psychology and the GPA which would be expected on account of his IQ. All subjects completed the Student Interest and Attitude Study II and the Minnesota Multiphasic Personality Inventory. The items which discriminated between the three groups were subjected to factor analysis and eight factors were identified. The results of this study were compared with a similar project carried out at the high school level, and the similarities and differences in outcome were discussed.

INTRODUCTION

A previous study of academic achievement at the high school level (Banreti-Fuchs, 1972) demonstrated a relationship between academic achievement and various attitudinal and situational variables, such as attitudes toward education, relationship to parents, attitudes toward authority and traditional social values, attitudes toward duty and responsibility, school behaviour and interests, and attitudes toward enjoyment and relaxation.

The results were essentially in agreement with other studies of variables related to academic achievement, such as Ringness (1967) on social conformity and rebelliousness, Khan (1969) on academic motivation, study habits, academic interest, and attitude toward teachers, Morrison (1969) on hostility toward authority, Oakland (1969) on work habits, motivational factors, and social conformity.

From the point of view of a more comprehensive theory of academic achievement, however, still many questions remain unanswered. We would have to know, for example, whether the factors which seem to be related to academic achievement at the high school level are the same as those governing academic achievement at the university level. In other words, are factors related to academic achievement stable across age levels, or do they vary as a function of age ? As far as variation of academic achievement as a function of subject matter is concerned, the study of Banreti-Fuchs (1972) clearly indicated that the achievement level of the high-, average-, and low-achieving groups did not vary appreciably across various subject areas. In view of this consistency of academic achievement across a variety of subject areas, the hypothesis is warranted that level of academic achievement is not a function of special abilities, but a function of more general traits affecting a person's efforts in a variety of academic subjects.

One of the objectives of this study was to examine whether the factors which were found to be related to level of academic achievement at the grade 9 high school level, would also be found to be related to the level of academic achievement at university. Another important objective was to find out to what extent mental health is a differential variable in level of academic achievement. Several studies have investigated the influence of anxiety on academic achievement, e.g., Davids, Sidman and Silverman (1968), Entwistle and Cunningham (1968) and Desiderato and Koskinen (1969), and have found that anxiety, in general, affected academic achievement negatively. Spielberger (1966) demonstrated that this conclusion was only true in the average range of intelligence, but was not applicable to people with very high or very low intelligence.

---

* Now at Department of Educational Psychology, University of Calgary, Alberta, Canada, T2N 1N4.

Although anxiety is an important aspect of mental health, especially as a major symptom of neurosis, it was felt that a more comprehensive measure of mental health should be used in an effort to determine its relationship to level of academic achievement. For this reason the MMPI was included in this study.

METHOD

The subjects in this study were 222 male and female students taking introductory psychology at Laurentian University. All students were administered the Henmon-Nelson Test of Mental Ability, Form A, and the Student Interest and Attitude Study II (SIAS II) in one session, and the Minnesota Multiphasic Personality Inventory in another session. At the end of the academic year when each subject's grade point average (GPA) in introductory psychology—based on three objective achievement tests—became available, all students were assigned to one of three groups : high-achievers (H), average-achievers (A), and low-achievers (L). Assignment to one of the three groups was based upon the regression equation method (Guilford, 1956) using the subjects' IQ scores on the Henmon-Nelson Test of Mental Ability and their grade point averages in introductory psychology ($r=0{\cdot}42$). Those students whose actual GPAs were between plus and minus 0·7 standard error of estimate from their expected GPAs were classified as average-achievers. Students whose actual GPAs were 0·7 standard error of estimate or more above their expected GPAs were classified as high-achievers, while students whose actual GPAs were 0·7 standard error of estimate below their expected GPAs were classified as low-achievers.

Item analysis was carried out on the responses of the three groups to each of the 135 items of the SIAS II and on the results of each MMPI scale expressed in terms of T-scores. Those items which discriminated between the three groups ($P<0{\cdot}05$ two-tailed) were intercorrelated and subjected to factor analysis using the principal component solution and the varimax rotation method (Kaiser, 1958). In deciding which variables to consider in the interpretation of the factors, a factor loading of $\pm 0{\cdot}40$ was chosen as the cut-off score, which was the same as in the previous study with high school students (Banreti-Fuchs, 1972).

In order to check whether the GPA in introductory psychology was an adequate criterion of academic achievement on the basis of which students can be meaningfully classified as high-, average-, or low-achievers, the grade point averages of the subjects in all the courses taken in that particular academic year were also computed. If the GPA in introductory psychology is a meaningful criterion of academic achievement, then the high-, average-, and low-achieving groups should show a similar distribution of achievement levels when the results of the complete academic programme of the students is taken into consideration.

RESULTS AND DISCUSSION

Since there were no significant sex differences in IQ at the various achievement levels, and the differences in GPA between the various achievement levels were equally significant for both males and females, the data of both sexes were pooled in order to increase N (Table 1).

TABLE 1

MEAN SCORES AND STANDARD DEVIATIONS OF DIFFERENT ACHIEVEMENT GROUPS.

Measure	Achievement Group			P<*
	High Mean (σ)	Average Mean (σ)	Low Mean (σ)	
Mean age	20·1 (2·1)	19·9 (1·8)	20·5 (1·9)	NS
Mean IQ	107·3 (11·5)	107·2 (12·2)	106·0 (9·7)	NS
GPA	75·0 (6·4)	60·6 (6·7)	46·1 (6·7)	0·01
(N)................	(55)	(119)	(48)	

* Using analysis of variance.

Seventeen items of the Student Interest and Attitude Study II and seven scales of the MMPI (F, Hs, D, Pd, Pt, Sc, Ma) discriminated significantly (P< ·05 two-tailed) between the high-, average-, and low-achieving groups. Factor analysis of the discriminating items resulted in eight factors (Table 2). There was no evidence of a strong general factor.

TABLE 2

ROTATED FACTOR MATRIX.

Variable	Factors								h²
	I	II	III	IV	V	VI	VII	VIII	
18	−·10	·08	*·79*	·12	·01	·07	·07	−·04	·67
19	·10	−·11	*·73*	−·24	−·00	−·22	·18	−·01	·69
24	−·10	*·63*	·06	·20	·23	·05	·11	−·01	·51
25	·01	·13	·05	−·02	*·74*	−·15	·07	·17	·63
33	−·09	·03	·09	·33	−·00	*·66*	·04	·27	·63
35	−·07	−·13	−·30	−·25	−·02	*·61*	·18	−·14	·60
48	·07	·02	·22	·01	−*·68*	−·03	−·10	·31	·61
69	−·01	*·40*	·03	−·16	·07	·12	*·61*	·10	·59
83	·02	−·22	·09	−·11	−·34	·31	−*·45*	−·23	·54
86	·04	*·79*	·14	−·07	·00	−·00	−·09	−·16	·68
93	·06	*·66*	−·24	−·09	·07	−·15	·12	·10	·55
95	·00	·00	·02	*·72*	−·00	·20	−·07	−·06	·56
106	·00	·12	−·20	−·19	·05	−·13	−*·66*	·13	·56
107	·01	−·02	−·06	*·69*	−·02	−·17	·18	−·04	·54
118	·04	·29	·27	·07	·34	−·04	·26	·39	·50
126	−·14	·22	·18	−·06	*·60*	·28	−·30	−·01	·63
135	−·00	−·12	−·10	−·12	−·05	·07	−·03	*·80*	·68
138	*·96*	·01	−·04	−·01	−·01	−·06	−·06	·01	·93
140	*·97*	−·01	·01	−·00	−·07	·02	·03	·01	·94
141	*·95*	·03	·07	−·02	−·01	−·06	·02	·05	·92
143	*·96*	−·03	−·01	−·01	−·05	−·04	—)01	−·04	·94
146	*·98*	·02	·03	·00	−·00	−·02	·02	·03	·96
147	*·98*	·03	·02	−·00	−·01	−·03	−·01	·01	·96
148	·94	−·05	−·10	·01	−·04	−·01	−·00	−·04	·90
Percentage of Variance	27·5	11·0	6·8	6·0	5·0	4·8	4·3	4·2	

Total variance accounted for : 69·6 per cent.

Factor I, after rotation, accounts for 27·5 per cent of the variance and could be labelled as a mental health factor. It is based on seven MMPI scales which discriminated significantly between the three achievement levels. The average T-score of the high-achievers comes closest to a T-score of 50, while the average T-score of the low-achievers is furthest beyond a T-score of 50, with the average-achievers somewhere in-between. If we accept the MMPI as an instrument measuring certain deviations from mental health, then the high-achieving group obviously comes considerably closer to the population average in mental health than the low-achievers.

Factor II accounts for 11 per cent of the variance and identifies attitudes toward duty and responsibility. High-achievers are more likely to attend lectures regularly, keep up with the reading material, finish assignments properly, and are more willing to put in time on studying even when they have the feeling that they are not getting anywhere.

Factor III accounts for 6·8 per cent of the variance and identifies self-concept regarding mental and physical health. High-achievers tend to rate their own mental health as better than average more often than low-achievers, while the situation with physical health is exactly reversed.

Factor IV accounts for 6·0 per cent of the variance and identifies attitudes toward social participation. High-achievers tend to have fewer friends and go to parties and dances less frequently than low-achievers.

Factor V accounts for 5·0 per cent of the variance and identifies adequacy of school adjustment. High-achievers tend to take notes more regularly when listening to lectures, they have been less of a behaviour problem in school, and the level of their academic achievement has been more stable over the years as compared to low-achievers.

Factor VI accounts for 4·8 per cent of the variance and identifies maturity and insight. High-achievers tend to be first-born children in greater proportion than low-achievers, and are able to predict their future academic performance more accurately than low-achievers.

Factor VII accounts for 4·3 per cent of the variance and identifies concentration of interest on academic achievement. High-achievers tend to keep up with the reading material better than low-achievers, they daydream less, and postpone marriage plans further into the future.

Factor VIII accounts for 4·2 per cent of the variance and identifies the need for self-expression in an examination situation. High-achievers tend to prefer subjective tests to objective tests.

When the overall Grade Point Averages of the high-, average-, and low-achieving groups were computed over all the courses taken during the academic year, the GPA of the high-achieving group was again highest, the GPA of the low-achieving group lowest, with the GPA of the average-achieving group in-between (Table 3). This fact demonstrates clearly that using the GPA earned in introductory psychology is quite an adequate criterion of academic achievement, even if one wants to make generalisations regarding academic achievement beyond the realm of psychology.

TABLE 3

OVERALL GRADE POINT AVERAGES OF DIFFERENT ACHIEVEMENT GROUPS BASED ON ALL COURSES TAKEN DURING THE ACADEMIC YEAR.

	Achievement Group			P<**
	High	Average	Low	
GPA	72·9	64·7	57·8	0·01
(N)*	(47)	(88)	(31)	

* Sample sizes differ as some students took only one course.
** Using analysis of variance.

Since achievement levels are defined here in relation to the subjects' measured intelligence, the relationship between the discriminating MMPI scores and each of the three achievement levels covers the whole range of intelligence represented in this study, covering an IQ range from 88 to 144. In other words, these results represent a clear indication of a relationship between academic achievement and measure of deviation from an average mental health level, unaffected by differences in IQ. The fact that the mental health factor (Factor I) accounts for 27·5 per cent of the variance in academic achievement between the high-, average-, and low-achieving groups, is an impressive indication of the role of mental health in academic achievement.

Some of the factors which discriminated between the high-, average-, and low-achieving groups at the grade 9 high school level (Banreti-Fuchs, 1972), seem to hold up well also at the undergraduate university level (Factors II, IV and V). Some of the differences in results between the two projects are no doubt due to the modification of several of the original SIAS items and the addition of new ones necessitated by the age differences between the high school and the university populations. The main differences in results between this study and the high school project are the following :

(1) The attitudes toward the parents, which formed an important factor in discriminating between the three achievement groups at the high school level, were not found to be important at the university level. This finding is not particularly surprising, since the role of the parents is expected to diminish rapidly during the years of late adolescence and early adulthood.

(2) At the high school level, high-achievers were characterised by greater conformity than low-achievers, and more receptive to authority and traditional social values. At the university level this is not the case. This change probably reflects the greater maturity of the university student who is more able and willing to judge for himself and has, in general, developed a much greater emotional and intellectual independence from parents, teachers, and all other representatives of authority.

(3) At the high school level, high-achievers preferred to concentrate on their work first, and to relax and enjoy themselves only after having finished their work. At the university level, this factor is also absent. This difference is probably due to the fact that the high school student has relatively short-term objectives, such as finishing homework for the next day. The university student, on the other hand, faces long-term objectives, such as deadlines for projects and term tests. He is much freer in deciding how he wishes to use his time. In his case, as Factor IV shows, it is the extent of social participation that is related to his achievement level, such as the frequency of his going to parties and dances, and the number of his close friends.

(4) The emergence of Factors III, VI and VII at the university level, which were not found at the high school level, may be ascribed to the differences in items between SIAS and SIAS II. This difference was necessitated by the fact that two different age groups were involved in the two studies. SIAS was geared to 14-year-old adolescents, SIAS II to 20-year-old students. Administering exactly the same items to both groups would have been obviously inappropriate. Although this difference between the two projects introduces an element of decreased comparability, it is an inevitable consequence of human development.

Acknowledgments—This project was supported by the Canada Council under Grant Number 572-0980.

REFERENCES

Banreti-Fuchs, K. M. (1972). Attitudinal and situational correlates of academic achievement in young adolescents. *Canadian J. behav. Sci.*, 4, 156-164.

Davids, A., Sidman, J., and Silverman, M. (1968). Tolerance of cognitive interference in underachieving and high achieving secondary school boys. *Psychol. in the Schools*, 5, 222-229.

Desiderato, O., and Koskinen, P. (1969). Anxiety, study habits and academic achievement. *J. couns. Psychol.*, 16, 162-165.

Entwistle ,N. J., and Cunningham, S. (1968). Neuroticism and school attainment. A linear relationship ? *Br. J. educ. Psychol.*, 38, 123-132.

Guilford, J. P. (1956). *Fundamental Statistics in Psychology and Education.* New York : McGraw-Hill.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187-200.

Khan, S. B. (1969). Affective correlates of academic achievement. *J. educ. Psychol.*, 60, 216-221.

Morrison, E. (1969). Underachievement among preadolescent boys considered in relationship to passive aggression. *J. educ. Psychol.*, 60, 168-173.

Oakland, J. A. (1969). Measurement of personality correlates of academic achievement in high school students. *J. couns. Psychol.*, 16, 452-457.

Ringness, T. A. (1967). Identification patterns, motivation and school achievement of bright junior high school boys. *J. educ. Psychol.*, 58, 93-102.

Spielberger, C. D. (1966). The effects of anxiety on complex learning and academic achievement. *Anxiety and Behaviour*. New York : Academic Press, 361-398.

(*Manuscript received 16th September*, 1974)

Br. J. educ. Psychol., 45, 232-236, 1975

PERSONALITY, MODE OF ASSESSMENT AND STUDENT ACHIEVEMENT

BY J. A. ROWELL AND V. J. RENNER
(*Department of Education, University of Adelaide, South Australia*)

SUMMARY. The EPI was administered to 136 full-time postgraduate Diploma in Education students. The relationships between personality, choice of method of assessment and achievement in four theory courses were investigated. Significant differences were found in each course favouring assessment totally by essay work ($P < \cdot 05$). In Theory of Education a significant interaction was found between neuroticism and assessment option, and in Educational Sociology extraverts achieved more highly than their relatively introverted peers. The hypotheses that introverts are more successful in structured courses and extraverts more so in relatively unstructured ones were tested and found to be supported.

INTRODUCTION

The Eysenck Personality Inventory (EPI) has been widely used as an instrument for obtaining information from subjects concerning two major, purportedly basic, unrelated aspects of their personality : extraversion and neuroticism. In turn, the possibility that such information is of value to educators has been examined in numerous studies covering a wide range of ages, abilities, courses and institutions. For those interested in prediction the results to date have not been particularly encouraging, however, as is evident in several recent summaries. Entwistle (1972) for example, summarises the situation by saying " It still requires considerable faith and imagination to see where these results may lead." Kline and Gale (1971) and Elliott (1972) are no more helpful. Only Leith (1973), who has utilised the technique of programmed instruction in an experimental attack on the possible relationship between personality and achievement in what he refers to as " school-type learning situations," gives reason for optimism. Leith maintains that introverts are more successful when they follow " . . . a carefully sequenced, highly prompted structure of learning;" extraverts, he claims, appear to learn better in less formal ' discovery ' situations.

The question is still a long way from any generally accepted solution ; meanwhile there is another dimension of the problem which appears to have received much less, if any, research effort. This extra dimension is mode of assessment. In any system where options are available the question arises naturally of the possible relationships between personality, option chosen and achievement. The research reported here addresses this question.

METHOD

Sample. 136 postgraduate students, 72 females (median age 22 years, range 20-41 years) and 64 males (median age 22½ years, range 20-40 years), representing 90·1 per cent of the 1973 enrolment of students to the one year full-time course leading to the Diploma in Education.

Test Instruments and Administration. All students and staff taking part in the investigation completed the Eysenck Personality Inventory, Form A, under the standard conditions described in the test manual (Eysenck and Eysenck, 1964). The majority of students for whom personality data were available completed the coursework assessments in psychology, sociology, history and theory of education.

Assessment of coursework provided students with a choice of seven options, the choice in each subject being independent of that in any other subject. The options allowed assessment by essay work only, by essay work and/or an oral plus a formal written examination each providing marks in various weighted proportions, or by formal written examination only. Because of the small number of students electing to be assessed by some methods, however, it was necessary for the purposes of analysis to collapse the seven categories into three. These were (i) essay work only, (ii) assessment in which the formal written examination counted for at least 80 per cent of the total marks, (iii) other assessment options.

RESULTS AND DISCUSSION

The mean extraversion (E), neuroticism (N) and lie (L) scores for the 123 students in the sample having L scores less than five are recorded in Table 1 by sex, separately, and as a group. The values are compared with the norms from the manual.

TABLE 1

MEAN SCORES AND STANDARD DEVIATIONS ON E.P.I. BY SEX.

Sample	N	Neuroticism	Extraversion	Lie
		Mean (σ)	Mean (σ)	Mean (σ)
Men students	56	10·59 (5·29)	12·18 (4·65)	2·07 (1·20)
Women students	67	12·49 (4·48)	10·67 (3·95)	2·00 (1·10)
All students	123	11·63 (4·94)	11·36 (4·33)	2·03 (1·14)
Eysenck norms	347	10·01 (5·01)	11·10 (4·54)	— —

In comparison with British undergraduates the Australian Diploma in Education students had higher N scores ($P < \cdot 01$, two-tailed) but there were no significant differences in E scores. Women students had, as expected, higher N scores than men ($P < \cdot 05$, two-tailed).

Personality and Achievement.

Chi-square tests were used to examine the null hypotheses of no differences between personality groupings in students' choice of method of assessment for each subject area considered separately. Because of the small number of students electing to be assessed by essay work only it was impossible to carry out such tests using the four EPI groups (NI, SI, NE, SE) and, instead, two tests were carried out for each course, separately categorising students as high and low N and high and low E, respectively—assignment to categories being determined by score relative to the mean of the group. None of the chi square values was significant. If choice of assessment is dependent on personality it was not apparent in this situation, given the limitations of measurement and analysis.

Two-way analyses of variance (2 levels of N or E ×3 modes of assessment) were used to examine the null hypotheses of no difference in achievement of students differing in N or E level, with the added dimension of comparative achievement between assessment options and the possibility of interactions. As mentioned above, numbers of students electing to be assessed by the essay method alone precluded use of the four EPI categories in these analyses.

Separate analyses were undertaken for each course, rather than treating scores as replications by individuals, principally because of the probable differences between courses in terms of the nature of the methodology employed in their teaching and the possible interaction between personality and methodology (see Leith, 1973 ; Entwistle, 1972).

Means and standard deviations of achievement scores, together with significant tests for the analyses of variance associated with each null hypothesis, are grouped in Table 2.

Three interesting results emerge from Table 2. Firstly, it is apparent that, in general, students choosing to be assessed by the option involving essay work alone achieved higher marks than students choosing other ways of being assessed. The option allowing assessment by essay work alone was the most demanding of student time and the corresponding investment of student concern and effort is one possible explanation of the higher marks.

TABLE 2

MEANS, STANDARD DEVIATIONS AND SIGNIFICANCE TESTS ON 2×3 ANALYSIS OF VARIANCE COMPARING PERSONALITY AND MODE OF ASSESSMENT.

Course	EPI Score	Ass. Mode	N	Mean (σ)	Hypotheses Tested	F	df	P<
Theory of Education	Low E	1	24	57·7 (5·0)				
		2	28	55·9 (4·3)	E Level	1	1,110	NS
		3	11	60·0 (5·8)	Assessment Option	4·36	2,110	·05
	High E	1	18	55·3 (5·0)				
		2	24	57·4 (4·8)	Interaction	1·66	2,110	NS
		3	11	60·5 (8·3)				
History of Education	Low E	1	22	61·0 (4·9)				
		2	32	59·3 (4·5)	E Level	1	1,109	NS
		3	9	62·4 (6·4)	Assessment Option	3·21	2,109	·05
	High E	1	16	58·7 (6·3)				
		2	27	58·7 (5·0)	Interaction	1	2,109	NS
		3	9	62·9 (7·9)				
Educational Psychology	Low E	1	16	58·1 (6·2)				
		2	37	58·1 (4·7)	E Level	2·02	1,113	NS
		3	10	60·0 (8·5)	Assessment Option	4·25	2,113	·05
	High E	1	16	53·9 (6·0)				
		2	26	55·7 (6·3)	Interaction	1·36	2,113	NS
		3	14	61·3 (7·5)				
Educational Sociology	Low E	1	16	54·4 (6·8)				
		2	37	56·2 (7·2)	E Level	6·22	1,114	·05
		3	11	58·1 (7·4)	Assessment Option	4·10	2,114	·05
	High E	1	19	56·7 (7·0)				
		2	24	58·6 (6·4)	Interaction	1	2,114	NS
		3	13	63·6 (7·1)				
Theory of Education	Low N	1	21	57·0 (5·7)				
		2	22	58·0 (3·6)	N Level	1·45	1,110	NS
		3	11	56·9 (7·1)	Assessment Option	4·35	2,110	·05
	High N	1	21	56·4 (4·5)				
		2	30	55·5 (4·9)	Interaction	6·48	2,110	·01
		3	11	63·6 (5·3)				
History of Education	Low N	1	19	59·2 (6·2)				
		2	27	60·1 (4·3)	N Level	1·11	1,109	NS
		3	7	64·6 (6·7)	Assessment Option	3·57	2,109	·05
	High N	1	19	60·8 (4·9)				
		2	32	58·0 (4·9)	Interaction	1·72	2,109	NS
		3	11	61·5 (7·2)				
Educational Psychology	Low N	1	15	55·9 (6·8)				
		2	27	58·0 (6·0)	N Level	1	1,113	NS
		3	16	61·4 (7·4)	Assessment Option	3·29	2,113	·05
	High N	1	17	56·1 (6·2)				
		2	36	56·5 (5·0)	Interaction	1	2,113	NS
		3	8	59·5 (8·7)				
Educational Sociology	Low N	1	16	54·9 (8·5)				
		2	29	59·1 (6·4)	N Level	2·22	1,114	NS
		3	13	62·8 (6·6)	Assessment Option	4·21	2,114	·05
	High N	1	19	56·3 (5·4)				
		2	32	55·3 (7·0)	Interaction	1·64	2,114	NS
		3	11	59·1 (8·5)				

Assessment options : 1 = 80 per cent marks or more from formal written examination.
2 = options other than 1 or 3.
3 = 100 per cent marks from essay work.

The second result of interest is the significant interaction (P< ·01) between N level and assessment option which occurred only in theory of education. High N students achieved much higher scores than low N students in the 100 per cent essay option and, indeed, much higher scores than high or low N students in any other option. Following Welford's (1965) ideas, a tentative theoretical explanation can be proposed. If N scores provide an indication of change in arousal level with stress, the opportunity to pace one's efforts and be assessed on a number of essays rather than a ' one shot ' end of year examination *may* provide the high N student with *one* of the conditions necessary for him to realise his potential—a relatively non-stressful situation. However, this explanation is essentially a tentative *post hoc* rationalisation unsupported by data from the other subject areas.

The third matter of significance is the fact that in educational sociology extraverted students achieved significantly higher scores (P< ·05) than their more introverted peers. The result is contrary to a great deal of other data concerning the relationship between academic success and extraversion (Furneaux, 1962 ; Warburton, 1962 ; Rushton, 1966 ; Lavin, 1967 ; Leith, 1969 ; Leith and Davis, 1969), although Killingsworth (1972) in an attempted replication of the work of Furneaux (1962), and Leith and Wisdom (1970) have reported similar findings. Leith (1973) has attempted to provide an explanation of his results in terms of an interactive relationship between method of instruction and the introversion/extraversion dimension of personality. According to Leith, " . . . extraverts have a greater tolerance for ambiguity and lack of structure in the teaching situation, whereas introverts are more inclined to be responsive to unambiguous and clearly structured situations."

Unfortunately, Leith's (1973) paper was seen too late to question the majority of students about the degree of structuring in the courses in educational psychology and sociology. (The history and theory courses had a common core of teaching and content.) However, there was unanimity among staff, and those students who could be questioned, that the psychology course had a higher degree of structure than the sociology. On the basis of this, admittedly limited, information two hypotheses were proposed for testing : that introverts would achieve more highly in educational psychology than sociology and that extraverts would achieve more highly in educational sociology than psychology. The results presented in Table 3 show both hypotheses to be supported.

TABLE 3

ACHIEVEMENT SCORES OF INTROVERTS AND EXTRAVERTS IN EDUCATIONAL PSYCHOLOGY AND SOCIOLOGY.

Personality	Educational Psychology Mean (σ)	Educational Sociology Mean (σ)	N	t	P<
Introverts	58·4 (5·7)	56·2 (7·1)	62	2·58	·01
Extraverts	56·6 (7·0)	59·1 (7·1)	57	2·83	·005

CONCLUSIONS

In 1962 Furneaux wrote, " It . . . seems to be clear that the results achieved by a student cannot be understood simply by considering his psychology, measured as it were in a vacuum. The characteristics of the university, and of its measuring devices must also be taken into account. It is only when the whole complex of student plus institution is studied that there can be any reasonable hope of obtaining useful information." Hopefully, however, we want to do more than passively ' understand ' why some have failed and others succeeded in a given institution utilising certain ' measuring devices.' We cannot alter the students ; one alternative is to alter at least some aspects of the institutions. A small contribution investigated here was the provision and utilisation of assessment options. The evidence presented indicates that students did not necessarily choose the option most suited to their talents ; in retrospect it is clear that they would not have the necessary information to do so. Students entering a one-year Diploma in Education course are confronted with a host of new variables ; new courses have to be coped with ; many mathematics and science graduates have little knowledge of their ability to write essays in any subject and, most important, there are new lecturers with initially unknown personalities, teaching methods, standards and expectations. The request by students to delay selection of their system of assessment is readily understandable in these terms, and justifiable too. They need time to establish criteria. The 1973 intake referred to in this paper were required to select their assessment option mid-way through term one, which may account for some of their difficulties. For the 1974 intake, time of selection of assessment option has been delayed until the beginning of the third term.

Finally, the results presented here suggest that EPI scores might be used to show how best to organise a set of parallel systems of teaching and assessing, in line with the course objectives, so that all students may have the *opportunity* to benefit to their fullest capacity. Leith's (1973) work has provided some testable hypotheses which have received tentative support in the data reported here. An ethically acceptable further step toward a ' broad tailoring ' of course presentation would be, with the permission of the students concerned, to identify personality types (NI, SI, NE, SE) and randomly to assign members of each in equal numbers to methods of instruction and assessment designed specifically to match their needs, as deduced from theoretical considerations, or to a common control pool taught by the ' usual procedures ' peculiar to the staff involved. Whether these procedures favoured one personality type would need to be assessed independently and predictions made accordingly. In this way relatively unbiased comparisons could be made between individual experimental groups and corresponding personality types within the control pool.

REFERENCES

ELLIOT, C. D. (1972). Personality factors and scholastic attainment. *Br. J. educ. Psychol.*, 42, 23-32.

ENTWISTLE, N. J. (1972). Personality and academic attainment. *Br. J. educ. Psychol.*, 42, 137-151.

EYSENCK, H. J., and EYSENCK, S. G. B. (1964). *Manual of the Eysenck Personality Inventory.* London : University of London Press.

FURNEAUX, W. D. (1962). The psychologist and the university. *Univ. Quart.*, 17, 33-47.

KILLINGSWORTH, J. S. (1972). Private communication to A. T. Welford. Reported in WELFORD, A. T. (1973).

KLINE, P., and GALE, A. (1971). Extraversion, neuroticism and performance in a psychology examination. *Br. J. educ. Psychol.*, 41, 90-94.

LAVIN, D. E. (1967). *The Prediction of Academic Performance.* New York : Wiley.

LEITH, G. O. M. (1969). Learning and personality. In DUNN, W. R., and HOLROYD, C. (Eds.). *Aspects of Educational Technology II.* London : Methuen.

LEITH, G. O. M. (1973). The effects of extraversion and methods of programmed instruction on achievement. *Educ. Res.*, 15, 150-153.

LEITH, G. O. M., and DAVIS, T. N. (1969). The influence of social reinforcement on achievement. *Educ. Res.*, 11, 132-137.

LEITH, G. O. M., and WISDOM, B. (1970). An investigation of the effects of error-making and personality on learning. *Prog. Learn. Educ. Technol.*, 7, 120-126.

RUSHTON, J. (1966). The relationships between personality characteristics and scholastic success in 11-year-old children. *Br. J. educ. Psychol.*, 36, 178-184.

WARBURTON, F. W. (1962). The measurement of personality III. *Educ. Res.*, 4, 193-206.

WELFORD, A. T. (1965). Stress and achievement. *Aust. J. Psychol.*, 17, 1-11.

WELFORD, A. T. (1973). Stress and performance. *Ergonomics*, 16, 567-580.

(*Manuscript received 29th October*, 1974)

BOOK REVIEWS

BRITISH COUNCIL FOR REHABILITATION OF THE DISABLED (1974). *People with Dyslexia.* London : pp. 161, obtainable only from the British Council for Rehabilitation of the Disabled, Tavistock House (South), Tavistock Square, London, WC1H 9LB, £2·00, plus 15p postage.

WADE, B., and WEDELL, K. (1974) (Eds.). *Spelling : Task and Learner.* Birmingham : Educational Review, Occasional Publications No. 5, pp. 68, obtainable from School of Education, University of Birmingham, 85p post free.

People with Dyslexia is the report of a working party set up in 1969 under the chairmanship of Dr. John Kershaw whose original terms of reference were to investigate the needs of the dyslexic adult. Oral and written evidence was taken by the working party whose members had between them both practical and research knowledge in a wide range of relevant areas. Questionnaires were also distributed to local education authorities, colleges of education, universities and examination boards to obtain information on treatment of dyslexics, teacher training and any special consideration given to dyslexic students.

Less attention has been paid to the important recommendations in the Tizard Report, *Children with Specific Reading Difficulties* (HMSO), with regard to early identification and appropriate remedial treatment and other facilities for all children with reading and related difficulties, than to its discussion of terminology. It is a matter of regret that there is, even now, so little evidence of a more sensitive awareness on the part of many psychologists, and those in authority, of the range of needs of children with severe reading difficulties ; and even, on the part of some, an air of complacency. This present report is timely, surveying, as it does, not only the needs of the adult dyslexic but also problems in the identification and treatment of children with severe reading, writing and spelling problems. The members of the working party realised very quickly that the dyslexic adult could not be considered in isolation and it was for this reason that they extended their terms of reference. This extension might well have resulted in further fruitless discussion centred on terminology or a discussion so wide-ranging as to be vague or superficial. On the contrary, the report is clear and well written ; dealing in a balanced unemotional way with the controversy. The reader is continually made aware of the distress and the frustration of 'the people' with whom the report is concerned. It is rare to find in an official report such a combination of sympathy and understanding and presentation of a variety of evidence all in an easy jargon-free style. This book should be read by all clinical and educational psychologists with responsibility for identifying children with special difficulties and for deciding the range of their needs ; also by those involved in planning courses for such children. There is a danger that an expansion of remedial education defined as withdrawal from the ordinary class from time to time in small groups, will be regarded as a sufficient provision for children with reading difficulties. The attempts at pressurising by some parents, conscious of their children's difficulties and frustrated by the lack of sympathy, or even sensitivity, on the part of many professionals, is commented on with understanding for the points of view of all those involved.

Spelling: Task and Learner contains the papers read to, and a summary of the discussions of a study group on spelling disability in children, held in the School of Education, University of Birmingham in 1973. One of the aims of the study group was to consider various aspects of spelling from different professional standpoints. As tends to happen with the proceedings of a conference, the papers are at a variety of levels and are unlikely all to be equally interesting to readers. Within the book it is possible to obtain insights into a number of aspects of spelling disability together with helpful suggestions on remedial assistance.

Although many of the contributors have other more lengthy publications on similar subjects, it is valuable to have in a single inexpensive book brief up-to-date reports from a group including Margaret Peters, Gill Cotterell, and David Moseley. As someone who attended the study course it is difficult to read the papers in a detached way and not to be influenced by the original presentation and a memory of the stimulating discussion they evoked. Much of the book is useful for psychologists in training and some also for teachers.

MARGARET M. CLARK.

CARPENTER, F. (1974). *The Skinner Primer : Behind Freedom and Dignity.* New York : The Free Press, Macmillan, pp. xvi+224, £3·95.

The empirical psychologist can deal best with the concept of freedom in very qualified, operational ways. Any psychologist, including the very eminent B. F. Skinner as well as those who comment upon Skinner's writings on freedom, should take this opening sentence very seriously. In issues about freedom, going beyond the operational meaningfully requires adroitness in philosophical inquiry at the least. Some familiarity with the immense philosophical literature on the concept of freedom and related issues should also help.

This mentioning of empirical psychology and philosophy is not offered in any zealous 'mind-guarding' between disciplines sense. It is simply that empirical psychologists are more likely than philosophers to research meaningfully operational conceptualisations of any behaviourally related construct, including those associated with freedom. On the other hand, philosophers are more likely to have the analytic skills and literary familiarity which will enable them to deal comprehensively with aspects of constructs which go beyond the operational.

Professor Carpenter's book is a courageous attempt to attend mainly to the psychology of freedom. The major virtue of this book is this reflected courage. The book's major weakness is its analytic imprecision. For example, Carpenter seems to be defining freedom by airing what he terms an assumption : " let us assume that freedom is the absence of anything that obstructs, blocks, or impedes a preferred or natural course of action " (p.122). The inclusion of 'natural action' in this definition is confusing. Many would argue that 'natural' actions are not free, obstructions not withstanding. That is to say, it is doubtful whether one is 'free' to sneeze or blink an eye whether or not obstructive conditions exist. Or is the reader expected to exclude autonomic reponses from the category of 'natural actions'? Other readers may be puzzled with this definition on other bases.

Such imprecision is also apparent in the fact that Professor Carpenter includes discussions which are not essential to an adequate psychological discussion of freedom. For example, he states that the essence of art is skill, and further, that art is any pattern of skill that is used to achieve a maximum effect (p.145). A broad definition of art could surely be developed without recourse to the intentions of artists. Another of these non-essential discussions in which imprecision is apparent is that of adolescent peer influences (p.139). The author claims that adolescents have an almost total inability to resist peer group influences. Published data do not support such a strong claim. Perhaps such imprecisions as these could be tolerated more readily if the discussions were essential to the main themes of the book. Since the discussions were not, one wonders why they were included at all.

The book would have been more robust both philosophically and psychologically had the space taken by such non-essential imprecisions been used to include discussions of more essential material. Discussions of some of the work of Kant or Sartre, to mention two philosophers, could well have been included. It is true that brief discussions of some ideas of Bohr, Bridgman, and Heisenberg were included and were directed primarily toward epistemological-scientific aspects of determinism rather than the metaphysical ; which seems to be consonant with Professor Carpenter's intentions for his book.

Even though this book is imprecise psychologically and operationally and even though it reflects a lack of familiarity on the part of the author with metaphysical literature, it is a laudable effort. Too few modern psychologists seem willing to render themselves vulnerable by working in areas which have such obvious connections with so much Western philosophy.

CHARLES K. WEST.

GEARHEART, B. R. (1973). *Learning Disabilities : Educational Strategies.* London : Henry Kimpton, pp. 233, £4·05.

This is an introductory text and it provides an overview of various systems and modes of provision for dealing with learning disabilities. The emphasis, inevitably perhaps, is American. It begins with a short historical review and a brief examination of various kinds of learning disability, concentrating on those which involve central nervous system dysfunction. There is a short chapter on evaluation procedures and eight chapters on major educational approaches. Chapters on behaviour modification, medical approaches and computer-based resource units lie uneasily together in a separate section. Four chapters are devoted to the structure and functioning of remedial services of various kinds and a final chapter provides a number of case studies.

A foreword assures us that Gearhart's approach is objective and non-judgmental : " . . . a reader can go through the entire book without feeling indoctrinated into any particular approach, system or point of view." This is not strictly true, however, for Gearhart's personal views on many issues are revealed in a number of ways. Reference to the 'psycho-motor *fraternity*,' for example (p.37), is clearly pejorative. Similarly, a reference to " eminent authorities " is curious in referring to such pioneers in the field of learning theory as Guthrie, Thorndike and Wertheimer, whose names are introduced as if to lend authority to a description of behaviour modification approaches, i.e., learning technology. The uncritical treatment of these approaches is, in fact, rather disturbing. In discussing Hommé's Reinforcement Menu, for example, there is no reference to the possible development of thoroughly undesirable learning sets, one of which could amount to, " do anything, however futile or foolish it may seem to you, if it is suggested to you by an authority figure who is prepared to reward you for doing it," These, and many other examples, tend to suggest that the ' unbiased description ' approach is really rather unsatisfactory. It would be less misleading, perhaps, for an author to state explicitly, if briefly, his general conceptual framework, relevant evaluative principles which inform his judgment on particular points, and his own position on major issues. The reader is then at least forewarned. Even more important perhaps, the novice is provided with a tool that may help him in learning to make his own independent evaluations.

In spite of these criticisms, the book could be of some value if each chapter is fully discussed, i.e., if the teacher does some of the job that might have been done by the author. The historical introduction is adequate, as is the preliminary examination of the nature of learning disabilities—although the introduction to problems of assessment is meagre. A useful selection of systems for dealing with learning disability is presented for inspection. The reader may not know what to make of these and they certainly do not enable a teacher to " formulate precise instructional plans " as promised in the foreword : nevertheless, these chapters do at least present a reasonable account of each system in varying degrees of detail. The descriptions of different kinds of remedial service are also interesting and useful—even though the reader is not invited to question the rationale of each form of provision. The individual case studies would be useful if linked to some kind of diagnostic-prescriptive activity—but such is not provided, or suggested. In short, this book contains useful material and, in the absence of anything better, could be used with profit by a teacher who is alive to its deficiencies.

JOHN E. MERRITT.

GEENSEN, M. (1974). *Retarded Children.* Groningen : Tilburg Univ. Press, pp. 133, Dfl. 29·90.

The Amsterdam investigation into the mentally deficient, conducted by the Institute for Social Research of the Tilburg School of Economics, Social Sciences and Law, under the supervision of Professor Godefroy, seeks to ascertain the number of mentally deficient persons, the relative incidence of the different grades of mental retardation, the need for care and guidance of the feeble-minded and the future need for institutions, sheltered workshops and other provisions in the city of Amsterdam. It is intended as a preliminary study to an extensive national survey.

The criteria laid down for this investigation define the mentally deficient as those who " do not fulfil the expectations that society has of its members and who score 79 or less in the intelligence test chosen for this investigation, provided that these conditions were already manifest at an early age," a definition that will no doubt cause a few eyebrows to be raised in this country. It must be seen as both tentative and provisional as it was to be critically tested for its usefulness during this pilot investigation.

The mentally deficient subjects traced *via* the services and institutions for the mentally retarded were given a sociological, a medical and a psychological examination. This book is a report of the first stage of the psychological examination, namely the examination of the 10- and 13-year-old children which took place during the period 1966-1967. The second stage, the examination of the adult subjects, followed in 1967-1969 and will presumably be published in due course.

Over 450 children were tested on the Vineland Social Maturity Scale, a Flemish version of the Wechsler Intelligence Scale for Children, and in the case of those children with a very low level of development, the Merrill Palmer Scale (Stutsman). The results of these tests are reproduced in detail in this report, followed by a comprehensive analysis including correlations and factor analysis of WISC scores, correlations between VSMS and WISC scores, the statistical regression theory and differences between Verbal Quotient and Performance Quotient (WISC) and Social Quotient and Intelligence Quotient, and sex differences in prevalence of mental retardation.

The author proposes answers to some important questions in mental retardation. Are the tests used in this investigation suitable for retardates ? Can the testing programme be shortened without loss of relevant information ? Are there differences between males and females in frequency of retardation ? Are there differences in the structure of intelligence between normal and mentally retarded subjects ? Can the large differences between Verbal and Performance Quotients and between Intelligence and Social Quotients be explained on the basis of regression theory ?

This is an informative and interesting report, particularly to psychologists, clinicians and researchers in the field of mental retardation. One might be critical of the tests used in the investigation with too much concentration on the cognitive aspects of development, but such criticisms are voiced by the researchers themselves. None of the tests was found to be appropriate to the severely retarded. Indeed, what tests are, among the many which are used with normal children ? Only by analysing the inadequacies of such tests when used with the severly defective can we hope to work towards tests that will be suitable, and this report certainly does that.

The reports of subsequent stages of the Amsterdam investigation and, one hopes, a combined report of the whole enterprise should add substantially to current research in this area.

A. HARDMAN.

LOWENFELD, B. (Ed.) (1974). *The Visually Handicapped Child in School.* London : Constable, pp. xvi+384, £3·50.

This essentially American text has much to offer the European audience concerned with the education, training and welfare of the visually handicapped. The wide range and experience of the contributors (10 of the foremost US specialists in the education of the visually handicapped) has resulted in an intensely *practical* book which is exemplified by a primarily down-to-earth tone which is refreshing.

The 10 chapters comprise a discussion of : the effects of visual impairments on cognitive functions and emotional social life ; understanding and meeting development needs ('development' covering physical, motor, emotional, social, personality, mental and educational development) ; psychological and educational assessment ; the utilisation of sensory-perceptual abilities ; educational programmes, communication skills ; special subject adjustment and skills ; life adjustment, visually handicapped children with additional problems ; and the preparation of teachers—which constitutes an impressive array of material.

However, the major strength of this text lies in its continual emphasis, and illustration, of the close similarity, rather than differences, of the visually handicapped population and the visually 'normal' population with respect to " the basic needs, rights, responsibilities and basic potential for fulfilment."

This is explicitly argued in Lowenfeld's succinct review of the *History of Education of Visually Handicapped Children*, which emphasises the transition from institutionalised separation and isolation to direct and normal community involvement. However, the reader should be aware that much of the US historical scene has not been paralleled in Europe.

Notwithstanding the philosophy of community-involvement, there is also emphasis (see Henderson, chapter 7) upon the four cardinal aspects of educational programmes for blind students, namely : (1) An enrichment of the core of communication by a wide variety of stimulating incentives ; (2) Provision of a realistic appraisal of the students' strengths and weaknesses ; (3) Sensible adjustment of teaching techniques ; (4) Development of a positive self-concept and self-motivation on the part of the visually handicapped student.

The book provides most extensive reference lists and bibliography.

Any weakness which the text contains may centre upon the extensive rather than intensive treatment of certain topics. But for those other than a few specialists this is acceptable, and even for the specialist there is much to be gained from alternative perspectives.

The text should prove to be an excellent source of material for future researchers and educators involved with the visually handicapped.

T. G. WHISTON.

MILGRAM, S. (1973). *Obedience to Authority.* London : Tavistock, pp. 224, £2·50.

So much has been written and spoken about this book in the popular media, some of it, indeed, has been excerpted, that a reviewer's style at this late date is somewhat cramped. We hope that the basic contents are more or less familiar to readers. To refresh overloaded memories, Milgram conducted a series of experiments in which subjects apparently gave shocks of gradually increasing intensity to an experimental stooge participating in a learning experiment. As the intensity increased the stooge cried out with pain, screamed, and finally ceased to respond at all. Variants were introduced into the experimental procedures, so that sometimes subjects had to hold the stooge's hand down on the electrical contacts, on some occasions there was the added danger that the stooge had a weak heart. Sometimes again the experimenter was not in the room with the subject but could communicate with him only by phone. The experimental measure derived from these experiments was the maximum intensity of shock that subjects were prepared to administer. Normally when subjects heard the first cry of pain, they would indicate to the experimenter that they should stop. However when the experimenter said one of four 'prods,' (" please continue," " the experiment requires that you continue," " it is absolutely essential that you continue," "you have no other choice, you *must* go on"), most subjects went on. This then, in brief, is the experimental paradigm of obedience to authority. Man's fatal flaw that has led to the Spanish Inquisition, the Nazi concentration camps and My Lai to select but a few from the variegated riches of Western civilisation.

If we accept the thesis that these experiments do reflect obedience to authority in the wider context of society, this book makes disturbing reading, for we see in it people who appear to be in no way pathological doing frightful things because they are told they must. However, there are some points about these experiments which we must discuss before we come to any conclusions.

In the first place we know that people are all too obedient, and our examples above tell us clearly enough that human beings can be inhuman. Thus, the fact that subjects gave severe shocks to people should not shock readers. The value of these experiments, then, lies in their capacity to tease out some of the conditions that enable people to do things that offend their conscience as individuals. The question about these experiments then becomes, "Do they enable us to do this?" One powerful objection to this work is the claim that subjects knew that they weren't really hurting the stooge. It was actually said in the experiment that the shocks caused no permanent tissue damage. Thus for subjects it was a case of trying to balance this statement against the contrary perceptual information of the screaming etc. Do we believe our ears or the experimenter? If we choose to believe the experimenter, then obedience to authority becomes hardly an issue. Certainly the study does not reflect, say, the concentration camps where we can see the effects of our actions. However, Milgram cites subjective evidence—of obvious conflict as the shocks were increased in intensity—and objective evidence in that where there was no supervision subjects cheated and gave low shocks, which indicates that subjects felt they were hurting the stooge. Thus it does seem that subjects were prepared to inflict pain for the sake of science. The introduction of a heart condition for the stooge hardly affected the results. Even if we accept that subjects felt that they were creating no permanent damage, it seems that they were prepared to inflict pain on request.

In fact, Milgram's programme of experiments neatly teases out some of the conditions affecting this obedience. Indeed, the book is an excellent model for the social scientist in that as each experiment raises a number of questions so the next experiment is devised to answer them. Furthermore, the style is refreshingly literate and free of jargon. The book, indeed, is concerned with the experimental analysis of a question thought to be important not just by psychologists (one buffer store or two) but by society as a whole.

It seems to this reviewer that, of experimental social psychology, this book is a typical example. It shows clearly just how ingenious experimentation can unravel a problem. It seems weak, however, in its treatment of subjects. What were the individual differences underlying differential performance? Within a psychometric framework, for example, what personality and motivational factors, as defined by Cattell, affected performance? Were ability factors important? Unfortunately we are told none of these things. Furthermore, Milgram's theory of an agentic state on some kind of cybernetic model seems trivial in the extreme—little more than a description in quasi-technical language.

To conclude: as a report of a series of thought-providing experiments this book should be read by all psychologists and, indeed, all interested in the problem of obedience. As a psychological work it is flawed by its incompleteness. It needed far more study of the subjects themselves and the theorising is, through this lack of information, necessarily weak. In this respect it compares poorly with the *Authoritarian Personality* or Rokeach's work on dogmatism. An excellent first half of a book is perhaps the fairest description.

P. KLINE.

MILLER, C. M. L., and PARLETT, M. (1974). *Up to the Mark : A Study of the Examination Game.* SRHE Monograph 21. London : Society for Research into Higher Education, pp. 128, £2·20.

THYNE, J. M. (1974). *Principles of Examining.* London : University of London Press, pp. ix+278, £1·85.

This review might be subtitled : Chalk and Cheese. The chalk is *Principles of Examining*, redolent of the lecture theatre, aimed at the teacher and examiner. The cheese is *Up to the Mark*, racy and appetising, aimed in the first instance at the academic staff and students of Edinburgh University and now in its published form at a wider audience, but still accessible to anyone interested in the debate about the uses and abuses of examinations.

Miller and Parlett studied the examinations of three faculties at Edinburgh University. They employed the strategy of investigation that has come to be known as 'illuminative evaluation,' which is eclectic in its methods. A substantial part of the text consists of quotations from interviews with staff and students ; questionnaires and non-participant observation were also used, but in all cases, the authors were careful to incorporate cross-checking and independent judgments of their classifications. The short sections and an appendix on the methodology provide a convincing defence of techniques that might appear, on the surface, to lack rigour, and amply demonstrate that the report is much more than a string of case studies, interspersed with the subjective interpretations of the authors.

The strength of their techniques is made especially clear in their original work on 'cue-consciousness.' They identified three types of student : two, the 'cue-conscious' and the 'cue-seekers.' were both very conscious of examination technique, but while the 'cue-conscious' were content to pick up hints from staff and to notice the particular academic interests of staff, the 'cue-seekers' actively tried to create a favourable impression on staff and button-holed them about examination questions (" This is another part of the examination game, you're always prodding. You know if you can get a lecturer over a cup of coffee, you are always saying ' Have you written the examination paper yet ?' you know, this sort of thing "). In contrast, the 'cue-deaf' would not believe there was an examination technique, or were unlikely to feel that the staff would be influenced by an impression of them when it came to the marking of their papers. They demonstrate the existence of a significant association between cue-consciousness and degree class, the 'cue-seekers' doing best and the 'cue-deaf' worst and also illustrate the way in which many staff condone the activities of the 'cue-seekers.' An examination game, indeed !

The coverage of the book is best illustrated by the five dimensions of assessment identified by the authors. These are task complexity (including reliability), the weight attached to the results (in their importance for the future) and consequent stress, task time, the predictability of the requirements of the task, and task distribution over time. Altogether this book provides a fascinating insight into the problems of university examinations, and the attitudes of staff and students. The postscript points to the importance of considering assessment methods in the context of the total circumstances—the ecological metaphor—and suggests a number of questions that flow from the metaphor. It is to be hoped that researchers will attempt to answer them.

Thyne's book, in contrast, is a textbook which appears curiously out of touch with modern examining methods and with current research. A large part of the book is concerned with theory, with the topics of validity and the weighting of component parts of an examination treated at considerable length. The author often seems more concerned with logical and semantic precision, enlivened occasionally by his own cartoons, than with practical advice to the intending examiner. One wonders what the uninitiated reader would make of this part of a sentence : "... because the measures obtained by measuring what we did measure would be equivalent to, and therefore as valid as, those that would have been obtained had we actually measured the performances we were supposed to measure."

Nevertheless, Thyne treats technical matters with competence and provides a useful discussion of norm-referenced and criterion-referenced testing. His advice on the preparation and marking of written questions (from true-false to open-ended) is sound, but the major weakness of the book is the omission of any discussion of other examining techniques such as oral and aural assessment, projects or dissertations, and continuous assessment. Ten years ago this book might have been useful to teachers ; now there are many very much better and more comprehensive ones available.

DESMOND L. NUTTALL.

OPEN UNIVERSITY (1974). An Introduction to Psychology ; *Block* 1, *Approaches and Methods* ; *Block* 2, *Psychometrics* ; *Block* 3, *Behaviourism* ; *Block* 4, *Information Processing Parts* 1 and 2. The Open University Press.

Judging by the continued demand for undergraduate places to read psychology it is a subject attracting ever-increasing interest amongst the young. The interest shown by 'mature' students has been evident for many years, and we may reasonably expect, therefore, that psychology will be in great demand at the Open University. Given the nature of the OU the number of students taking their psychology courses will probably be very large over a five-year period. From this we might expect that their course units, books and general orientation will have a considerable formative influence on a large proportion of British psychologists of the future. How good will this influence be ?

If the first five volumes, units 1-13, are a reliable guide we have little to worry about. This second level course in the social sciences is refreshingly different from most of the available introductory texts. The first three units, *Approaches and Methods* lay out the scope of the subject in very broad terms and relate it to other disciplines. Psychoanalysis and introspectionism receive a surprising amount of attention but this may be because their subsequent treatment is to be fairly slight. The third section of this volume produces a very readable and elementary introduction to statistical ideas. This would be an excellent volume for sixth-formers to read before embarking on a degree course.

The second volume, units 4-6, is surprisingly entitled, *Psychometrics*. Surprising not because of its content which is very clear but because it is brought into the course, in quite a detailed way, at this early stage. The theoretical basis of testing is discussed along with the procedures involved in test construction. The use of psychological tests in various settings is dealt with thoroughly and sensitively. Given the controversies which presently rage, often mindlessly, about psychological testing, the short section on ethical problems is a very sensible and level-headed inclusion. The final unit in this volume is concerned with personality and individual differences. This is, perhaps, the weakest section of the book though it does bring in factor analysis in a way easily understood. I am delighted to see this volume—it will provide a very welcome alternative to the dry introductory books on testing presently available for undergraduates.

The third volume, units 7-9, came as something of a surprise. It is entitled simply, *Behaviourism*. The first unit is very largely an historical appreciation of the roots of behaviourism. It reads to me like an elegant television lecture which might well have been delivered by a 'psychological A. J. P. Taylor' ; it needs supplementing by further reading. The second section provides an excellent introduction to learning. Basic concepts in animal and verbal learning are simply outlined. The final section deals with the applications of behaviourist research and a critique of it. Much of this is set in a clinical context and surprisingly little is said of educational technology. I suspect the anti-behaviourist school would wish for more detailed and more trenchant criticism.

Units 10-13 deal in two volumes with human information processing and bear that title. They cover, in an impressive way, feed-back, goal-directed behaviour, perception, memory and skills. They represent what in the jargon of the '70s is known as cognitive psychology. A striking feature is their difficulty level in comparison with the preceding units. They are clearly written, cover a great deal of ground and do a very good job presenting tough, quantitative material in a lucid and streamlined fashion. Inevitably in a collection covering such an enormous range of subject matter as these volumes the reviewers will find some sections more to his liking than others. The sections on feedback and memory seemed particularly good. These two volumes provide a welcome addition to the literature suitable for undergraduates in the early part of their course.

Any summary evaluation of OU material is difficult but given the objectives of this collection, as I understand them, it succeeds very well. The material is presented vividly and the design of the books is excellent. Indeed, textbook publishers would do well to study them. The subject matter covered is very carefully and effectively distilled so that the student has a clear picture of the elements of the question, for example, information theory and reaction time studies in unit 13. Questions, experiments, references and a smaller sampler of journal articles are all included. Given that the students supplement these texts by wider reading, in which more detailed consideration is given to theories and methods, they form an excellent launching pad.

KEVIN CONNOLLY.

POLLIO, H. R. (1974). *The Psychology of Symbolic Activity.* Reading, Mass. : Addison-Wesley, pp. 504, n.p.

Human thought processes are marvellously diverse, wide-ranging and ingenious, and these attributes are undoubtedly present in the approaches by which psychologists have investigated reasoning and symbolic activities. It takes a brave author to produce a book that brings together the findings of the confusingly varied investigations that have been undertaken, and a capable one to impose any real order upon the scattered bits and pieces of information that the findings comprise. It would require a miracle-worker to form these bits and pieces into a coherent and integrated body of knowledge, coming as they do from the efforts of individuals attacking the diverse problems by a variety of methods, and holding a multiplicity of opinions concerning the relative efficacy of alternative approaches to the study of human reason. Fortunately, Professor Pollio is both brave and able, and he has produced a lucid and interesting survey. Unfortunately he can perform no miracles, and the research he surveys remains a confusing assembly of incompletely integrated insights and pieces of knowledge.

The book is divided into four parts. The aim of part one is to describe some of the ways in which symbolic activities enter into human life, and to explain some of the complexities, in order to set the stage for the psychological analysis which comprises the remainder of the book. Pollio points out that novelists have often been more successful than psychologists at describing human thought processes, largely because they are not subject to the constraints required for scientific acceptability. Taking two prose passages, one from *Swann's Way* and the other from *The Old Man and the Sea*, he illustrates the snares involved in investigating thinking, and examines the manner in which psychologists from alternative traditions, associative and information processing, might attempt to describe and throw light upon the data.

The distinction between associative and information-processing approaches to thinking forms the dividing line between parts two and three. Part two considers memory and forgetting, concept formation and the development of thought, within an associative framework. Ostensibly, similar topics are encountered in part three, within the context of the information processing kinds of approach favoured by modern structuralists. Thus, while part two contains chapters headed 'Concept Formation,' 'Problem Solving' and 'The Development of Thought,' virtually identical chapter-titles appear in part three, namely 'Forming and Using Concepts', 'Solving

Problems' and 'The Development of Thinking.' The fact that Pollio found it necessary to make the broad structuring of his book follow the alternative approaches to his subject matter provides an indication of the restricted extent to which the findings of the different brands of psychology have been mutually contributory. As it happens, Pollio avoids making a judgment concerning the relative values of the different approaches. After noting that associationist and information-processing views have led to different traditions of theory and experiment, with little overlap, he states (p.294):

> "The strange fact is that both views have been able to generate reasonable and extensive research programs which make sense not only to proponents of a particular view, but also to anyone who has a less predjudiced stake in the matter. There are times when the mind functions as if it were guided by perfectly rational plans, and there are other times when it seems to function strictly on the basis of contiguity."

Part four contains three lengthy chapters on language. After providing a formal analysis of language Pollio turns to the problems of the user in producing and receiving human language. Once more the contributions of researchers in the associationist tradition are presented in a separate chapter, followed by a chapter which considers the contributions of researchers following information-processing approaches to knowledge about the use and development of language in the individual.

Professor Pollio writes with enviable clarity, and throughout his book he is at pains to emphasise the recent and ongoing changes in our conception of human language, thinking and problem-solving. Pollio is very much concerned with the ever-changing nature of the scientific activities by which we enquire into symbolic activities. He has produced an excellent text which provides a readable tour of those areas of human symbolic activity on which psychological investigation has shed some light.

MICHAEL J. A. HOWE.

ROUTH, D. K. (Ed.) (1974). *The Experimental Psychology of Mental Retardation.* London : Crosby, Lockwood Staples, pp. 353, £6·00.

The editor of this welcome publication is a Clinical Associate Professor of Psychology at the University of North Carolina in the USA. There are seven other psychologists, including Dr. N. O'Connor from London, who contribute eight chapters of varying length which adds to about 350 pages in all.

The purpose of this book is to remind the workers in the field of mental deficiency that until quite recently this subject was a neglected field of study. The well-known book *Mental Deficiency : The Changing Outlook* by Clarke and Clarke appeared in this country in 1958. Similarly on the other side of the Atlantic the monograph by Masland, Sarason and Gladwin entitled *Mental Subnormality : Biological, Psychological and Cultural Factors* was also published in the same year. There were several other publications and practical actions on behalf of the mentally deficient in the 1950s. Looking back, it appears ,therefore, that the period since World War 2 can justifiably be described as the renaissance in the area of mental retardation. The present book and the symposium on which it is based attempt to continue this tradition by presenting programmatic research in the field under discussion, aimed at the scientific understanding of the psychological processes involved.

It is suggested in the book that the retarded child differs in systematic ways from the normal child, and although the ensuing research is conducted more in the laboratory than in the consulting rooms, its results are ultimately and directly relevant to understanding and treating developmental disorders. The contributors here compare retardates to normals, to each other and to themselves, using many research examples to describe and extend new ways of thinking about these children. For example, N. O'Connor focuses on new ways of testing the ability of the retardates to discriminate and Leonard and Susan Ross explain conditioning and reinforcement which can be used to help us understand the learning processes of the retarded and to reflect their basic capabilities.

David Zeaman discusses discriminative learning, Norman Milgram discusses cognition and language and the developmental lag of the retarded, and Edward Zigler studies the retarded child as a whole. The editor of this volume contributes a summary essay on the experimental study of the retarded child and a concluding chapter describing in general terms the several experimental approaches to the clinical psychology of mental retardation. The conclusions of the book are important not only to the psychologists in the field of mental retardation but also to all those who are striving to create positive changes in the handling and treatment of retarded children.

VED VARMA.

RUBIN, L. J. (Ed.) (1974). *Facts and Feelings in the Classroom*. London : Ward Lock Educational, pp. 287, p. £1·95, c. £3·50.

Why do I so often get to review books I don't think very much of ? Is it me or is it them ? Here we have a well-intentioned attempt to deal with some important problems, by a variety of authors including such distinguished names as Michael Scriven and Abraham Maslow. But I have to say that I find it very unsatisfactory.

The book has two main messages. One is that school education has neglected the emotions in favour of the intellect, and that we should redress this balance. The other, closely related, is that the aim of education should be to make children (and teachers) more autonomous. The programme is deliberately one of moderate reform. This seems unexceptionable : to what, then, do I object ? What worries me is that I cannot really find anywhere in the book anything that tells us *why* we should have such a programme, nor yet *how* we should put it into action. For the most part, the authors deal in generalisations and assertions, with little hard evidence in support. There may well be such evidence ; but we do not find it here. Where research findings are reported, they lack information as to procedure and context, and seem often of paralysing obviousness.

"Bloom discovered, for example, that 'unsuccessful experiences in school guarantee that the individual will develop a negative academic self-concept and increase the probabilities that he will have a generally negative self-concept." What a breakthrough !

"In an imaginative departure from conventional research technique, Bloom used an ideal tutor as his model of good teaching. From this model, he deduced that effective teachers make abundant use of cues that clarify the learning task for the student ; they provide the learner with many opportunities to practise the responses to be learned ; and they use a variety of reinforcements, both positive and negative, as well as both intrinsic and extrinsic." Apart from the difficulty of whether one can deduce conclusions from a model, there can have been few good teachers, from Socrates onwards, to whom these remarks would come as much of a surprise.

I forget who it was who said that the secret title of every book ever written is "How to be more like me." In this case "me" is presumably Louis J. Rubin, who has edited the book, contributed four of the 11 chapters, and commented on each of the others, written by Ralph Tyler, Edward Meade, Michael Scriven, Benjamin Bloom, Abraham Maslow, Richard Jones and Elliot Eisner. Much of what they have to say can be found elsewhere. This is not a criticism since for many readers it may be useful to have a concise compilation of views. Probably two sorts of reader are in prospect : the interested layman and the educationist. For both the book may serve a purpose in raising questions about what is, after all, our most peculiar institution of universal compulsory education. But I hope that readers will also have access to rather more sophisticated arguments, and firmer evidence, than are presented here. I would also suggest that the issues might be raised more interestingly and illuminatingly by other writings, such as those of John Holt, Paul Goodman, or Alex Comfort or even, perhaps, Froebel, Rousseau or Plato.

JOHN RADFORD.

SCHWARTZ, L. L. (1972). *Educational Psychology : Focus on the Learner. Boston* : Holbrook Press, pp. x+604, £11·50.

Satisfying textbooks of educational psychology seem to be extraordinarily difficult to write. This probably stems partly from the nature of the subject itself—it embodies a variety of forms of explanation which are often confounded and confused—and partly from the sense of obligation to provide something of relevance to prospective and practising teachers, when most of the problems that teachers have to cope with are not the kinds of problem that educational psychology is equipped to tackle. Faced with this dilemma, the textbook writers tend to veer in one of two directions ; either they produce a book divided up into the conventional headings of ' personality,' ' motivation,' ' learning ' and the like, or they make a determined attempt to concern themselves with learning in schools. The weakness of the former is that they tend to be full of flat empirical generalisations useful for passing educational psychology examinations, but only tangentially related to what goes on in the classroom ; the latter tend to suffer from trying to validate in psychological terms what is already held to be true on commonsense grounds. *Educational Psychology : Focus on the Learner* is more of the second kind, although the familiar areas of educational psychology are represented in somewhat modified form.

It is a nicely produced and simply written book which, in its own terms, succeeds very well. It first tries to develop a sense of perspective by examining the assumptions and outlining the techniques of educational psychology, and then it concentrates, in turn, on the learner, learning, and the teacher. Consideration of the learner includes chapters on growth and development, environmental influences, and personal and intellectual factors ; learning is examined in the context of its theories, content and methodology ; and the teacher is viewed as a learner himself. Thus, what we are offered is essentially the conventional wisdom of educational psychology specifically interpreted for teachers, with attempts to tease out the implications for classroom practice.

The book is well set out and very easy to find one's way around. Each chapter begins with a preview and ends with a summary, and is supported by a glossary and bibliography. A special feature is the " Focus " which picks out particular issues and provides material to stimulate reflection and discussion. In the main these are well chosen (for example, a somewhat optimistic view of educational psychology is balanced by J. M. Stephen's critique) but they do flag at times. Thus, a wholly worthwhile attempt to expose the banality of some teacher evaluation schedules by rating Socrates (he comes out as "*Does not have a place in Education. Should not be hired*") loses its impact because of the forced and inaccurate answers to some of the items.

However, I enjoyed reading *Educational Psychology : Focus on the Learner* more than most educational psychology textbooks. Clearly a lot of thought has gone into its content and presentation. I should imagine that it will have a wide sale in America, but be too American to replace the established texts in this country. If at times it is trivial, then I think this is because it reflects the triviality of much of what is, at present, accepted as educational psychology. It may be that we ought to begin to think seriously about trying to develop a practical organisation of knowledge about teaching. This would probably be normative, be drawn out of teachers' own generalisations on how to cope with their situation, and leave educational psychology to grow to meet its tasks of description, illumination and explanation.

ALAN SMITHERS.

SPODEK, B. (Ed.) (1973). *Early Childhood Education.* Englewood Cliffs: Prentice Hall, pp. 280, £2·50.

According to the editor and main author, the origin of this book lies in the wide expansion, both in quantity and variety, of early childhood education in the United States over the past decade, and the corresponding need for teachers, parents and others in education to be able to make informed choices among the various educational programmes for young children now being offered. Spodek states that there is an increasing tendency for parents to be consulted by educationists about the kind of programme they want for their children, and to be offered choices among alternative proposals " much as an architect might present drawings and plans to a client for whom he is constructing a building." This book, therefore, aims to provide a concise guide to the historical background of early childhood education, the major issues in this area, and the theoretical and philosophical models underlying modern educational programmes. The term ' early childhood ' covers roughly our pre-school and infant school stages.

The book is divided into two parts. Part one comprises four chapters, written by Spodek, which provide a general background to the description of particular programmes presented in part two. In the first chapter the author examines the case for institutionalised early education, citing social, moral and psychological justifications. The second chapter provides a conceptual framework for the analysis and appraisal of educational programmes in terms of their assumptions, goals, curricula, methods, styles and organisation. Evidence for the effectiveness and practicability of each programme should also be sought but, as Spodek states, this is not always available. Chapter three discusses some historical models of early childhood education, including those of Froebel and Montessori, using the above framework.

The fourth chapter in this section contrasts two opposing contemporary psychological models of human behaviour, the behaviourist and the phenomenological, and briefly describes their application to the education of young children. Spodek claims that current educational programmes can be placed on a continuum with these two models at opposite ends. Thus the 'ideal behaviourist' model would be characterised by a statement of programme goals in behavioural terms, a very high degree of structuring and teacher direction of learning experiences, and the use of ' extrinsic ' rewards (at least initially). The approach is a ' closed ' one in the sense that the child's responses are pre-determined. In contrast, the ideal phenomenological or ' open education ' model is child-centred in the sense that the child is given considerable responsibility for selecting learning experiences and is encouraged to interpret them in terms of his own conceptual system. ' Intrinsic ' motivation is stressed and the aims and structure of the curriculum tend to be implicit. The programme developed by Bereiter and Engelmann exemplifies the former model, whilst the British primary school, at least in its more progressive form, illustrates the latter.

Part two of the book contains some further chapters concerning the historical background, assumptions and issues in early childhood education, but mainly comprises papers describing a wide variety of papers. These are by various American authors but, unlike the chapters by Spodek in part one, are not original, being drawn from various educational and other journals. As well as programmes representing the more extreme positions on Spodek's continuum there are those occupying a more central position. For example, the ' Piaget Derived Pre-School Curriculum ' described by Kamii has very explicit cognitive goals but the child is guided toward achieving these in open-ended rather than highly structured learning situations. Whilst the emphasis is upon teaching culturally disadvantaged children, some of the papers discuss the educational needs of children more generally and one writer at least examines the needs of *advantaged* (middle-class) children.

Several of the programmes described have clear similarities in approach and content to the British infant school but tend to be distinguished from the latter by their sheer scope and degree of systematisation. For example, they often include a consultation service and training scheme for teachers, planned involvement of parents and built-in provision for the evaluations of programme effectiveness. In one paper the virtues of the modern British primary school are extolled at the expense of the American elementary school system which is characterised as didactic and formal. It is interesting in this connection, however, that, at least as far as the education of disadvantaged children in Britain is concerned, there is an increasing emphasis on the need to introduce more structure and teacher direction into the learning situations provided.

In summary, this is a concise but readable book which provides a useful outline for parents, teachers and educationists of the historical, philosophical and psychological background to early childhood education and of the range of current approaches to the teaching of young children. It is refreshing to note that the British contribution in this sector of education is well represented. The major author's proposed framework for the appraisal of educational programmes is helpful but one would need considerably more information than is provided in the various programme descriptions, in order to make choices among them. From this point of view part two of the book would have been more useful if these descriptions had been specifically written for the book, using Spodek's framework. Spodek's own contributions are sound and fair-minded.

T. Cox.

WILKINSON, A., STRATTA, L., and DUDLEY, P. (1974). *The Quality of Listening*. London : Macmillan, Schools Council Research Studies, pp. 111, £3·00.

Oral assessment, particularly that variety experienced by candidates in CSE English examinations, has conventionally been the assessment of speaking, rather than listening, abilities. The Oracy Research Unit based at the University of Birmingham School of Education was sponsored from 1966 to 1972 by the Schools Council to investigate the whole field of speaking and listening, including its assessment. In this report from the unit the authors detail the outcome of the operational focus of one of its aims : the construction of tests of listening abilities appropriate to the 11-18 age range. Three batteries of ' Listening Comprehension Tests ' (A,B, and C) have been produced for the 10+, 13+ and 17+ age groups respectively. The characteristics of spoken language selected to be the features tested in separate test items were the cognitive content of speech, the constraints imposed upon language by the context of the utterance, phonology, register and between-speaker relationship evident from language use. Tapes were produced aiming to simulate appropriate speech situations in which all the various selected elements could be tested in turn, the testees responding to the tapes at specified points by marking answers in an answer book.

The form of the report is by no means confined to an account of the construction of the tests, however. The book is patently designed to appeal to a wide readership, to stimulate interest in what the writers call the most neglected of the four language arts—listening. Chapters 1 and 2 contain brief but clear expositions of the various channels of human communication and their interactive nature. Language code, lexis, grammar, phonology and register, together with a neatly drawn distinction between spoken and written language, are briefly discussed. In Chapters 3 and 4 the features of language selected for the construction of the tests are related more specifically to examples from the tests themselves and from the explanatory taped materials accompanying the tests. The reader is thus given an indication of the rationale informing the construction of the tests and the flavour of their form. Chapter 5 enlarges upon reasons for poor listening, and the writers attempt to indicate how improvements in listening ability can be achieved in the classroom by arranging for tapes of speech to be played to students, who would subsequently discuss their impressions of the taped speakers. The final chapter, of possibly greatest appeal to the research worker, is devoted to the detailed account of the construction and trials of each battery of tests, together with assessment of the measured relationship between performance in the tests and scores on intelligence, reading, and personality tests. The writers admit that the validity of the tests is open to question, since no criteria were available for validation, except for the ' face ' validity produced by consensus amongst a group of informed people.

As a means of promoting greater interest in listening and listening abilities amongst teachers, teachers in training and ' intelligent laymen ' (Preface), the book should be of value. The comprehensive, but concise, discussion and analysis of spoken language already referred to are non-technical, explicit and useful introductions to the subject. The weakness of the book, as distinct from the tests it partly describes, seems to be an odd inconsistency between the substance of the research it reports and the functions that the writers recommend the completed test batteries might serve. For example, the writers are not happy about the ' training of skills ' approach as applied to listening, and comment adversely upon it, because it seems to them to be a naive way of trying to externalise and describe a ' complex internal process ' (p.65). They urge that any attempt to improve listening should begin from the standpoint that any improvement is interrelated with other aspects of oracy, and even with general development. Yet the form of the tests described in the report could, if they became popular, tend to encourage just such attitudes and practices which are based on characterising ' listening ' as a set of discriminable and testable skills. The writers suggest that the tests could be used diagnostically, but are less than enthusiastic about having them used as models for external examinations. A stronger claim made is that the tests will have a beneficial effect on classroom practice as they will promote an awareness of ' the features and functions of spoken language ' (p.65) A further claim is that, through knowledge of the tests, teachers in training will be provided with knowledge of the spoken language important to all teachers. The writers are, therefore, claiming major teaching (rather than testing) functions and information-giving (rather than information-demanding) functions for these tests.

Having made these perfectly proper claims, they do not make clear whether the teacher's increased knowledge is seen as a step along the way to (*a*) imparting knowledge to pupils as knowledge *about* language, or to (*b*) using that knowledge in such a way that a pupil's listening abilities will be improved, or to (*c*) using that knowledge so that a pupil's subsequent performance on the tests would be improved, none of which is a necessary condition for both, or either, of the other two.

It is noticeable that the Schools Council's foreword to the report contains the disclaimer that the presentation of the report implies that the Schools Council is advocating large-scale-scale testing of listening comprehension. As the research began in 1967 with the aim of providing precise measures of listening ability in a given age range, it does seem odd that the assessment function of the tests is minimised whilst other applications, such as increasing the awareness of spoken language in teachers, are stressed.

This aside, the unique and exploratory nature of the work reported will be of interest to many. Those interested in measuring or improving listening abilities, or in language competence in general, will find material in the report to provoke thought about the forms oral English in the classroom could and should take.

H. MERCER.

WITKIN, R. W. (1974). *The Intelligence of Feeling*. London : Heinemann, pp. 198, £3·75.

In a book intended for a ' diverse audience ' including psychologists, sociologists, educators and teachers it is almost inevitable that some will feel hard done by. Witkin's work is thought-provoking and challenging, and, simply, too diverse. It includes an epistemological analysis of the creative arts, a probing of educational assumptions, and a report on fieldwork carried out in a number of schools—all this built round a core based on seeing artistic expression as a dialectic between impulse and medium.

The first part of the book lays the epistemological foundations of Witkin's approach, and is, in many ways, intriguing and yet unsatisfactory. Despite a subtle framework, based on the dichotomies of subject and object, impression and expression, and reaction and reflection, Witkin makes the world too simple. Science, logic and mathematics are written off as " relatively pure media of impression," and creative tension is simplified to oscillation between the sensate impulse and the medium of expression. He would have done well to draw on recent psychological literature on creativity outside the arts, and, even more important, an existential philosophy: not only is Sartre missing from the general bibliography, but so is Merleau-Ponty, whose recent *Prose of the World* is so pertinent to the point.

The second part of the book, an account of data collected in a Schools Council project, has its own problems, too. Here there is an annoying lack of data about sample sizes, questionnaires, interviews, the coding of classroom interaction, and so on. An appendix would have been helpful, giving all the usual background data in the study. These grumbles having been made, however, it is in this part that the book really takes off, and Witkin's account of the concerns and criticisms of teachers of drama, literature, music and art are fascinating and telling. I was particularly taken with his analysis of the conflict many teachers feel about releasing and yet containing their pupils' emotions, and the complexities they face in trying to convey technique, and yet allow self-expression and innovation. He also makes telling points about the effect of institutional demands, especially assessment. The analysis of the pupils' responses to the questionnaire is less satisfactory, however, and many questions are left unanswered: as an example, a question about an 'ideal' school timetable reveals a swing to vocational subjects and physical education, which Witkin interprets as evidence of a desire for personal involvement. He misses the more obvious interpretation that the pupils prefer these as they are the less academically demanding, non-examinable parts of the curriculum, an interpretation reinforced by his own data on the increase in this tendency as the more 'academically' inclined schools are considered.

The final part of the book is a brief statement of Witkin's own views on the changes that should take place. Sadly, though clear and reasonable, it stands rather lamely at the end, and needs a much fuller statement. Even better would have been, an attempt to try out his ideas, and a commentary on the consequences. Instead, the final paragraph is a call to freedom, a rally to the banner, that feels out of place. The book has shifted once more, from theory, through analysis and prescription, to end in rhetoric.

Criticisms having been made, this is nonetheless a valuable book. Witkin has pointed up some of the contradictions and controversies that bedevil the teaching of the creative arts in a way that is illuminative and challenging. Out of his diverse audience, I hope it finds its way into the hands of teachers in particular, where its effects could, and should, be far-reaching.

PETER SHELDRAKE.

PUBLICATIONS RECEIVED

Books for review have been selected from the following. (" n.p." means " no price stated.")

ABBS, P. (1974). *Autobiography in Education.* London : Heinemann Educational, pp. 182, p. £1·25, c. £3·50.

APPLETON, I. (Ed.) (1975). *Leisure, Research and Policy.* London : Scottish Academic Press, pp. xiv+255, £4·00.

BARKER, P. (Ed.) (1974). *The Residential Psychiatric Treatment of Children.* London : Granada, pp. xiii+353, £7·50.

BENNETT, W. A. (1975). *Applied Linguistics and Language Learning.* London : Hutchinson Educational, pp. ix+352, £5·75.

BERGER, N. (1974). *Rights : A Handbook for People Under Age.* Harmondsworth : Penguin, pp. 160, £0·60.

BILLER, H. B. (1975). *Paternal Deprivation.* London : Lexington, pp. 226, £6·85.

BION, W. R. (1975). *Attention and Interpretation.* London : Tavistock, pp. 136, £0·85.

BLACKMAN, H. J. (Ed.) (1975). *Ethical Standards in Counselling.* London : Bedford Square Press, pp. 69, £0·80.

BRENNAN, W. K. (1974). *Shaping the Education of Slow Learners.* London : Routledge, pp. x+114, p. £1·25, c. £2·50.

BROWN, G., CHERRINGTON, D., and COHEN, L. (1975). *Experiments in the Social Sciences.* London : Harper and Row, pp. 256, £2·95.

BRUNER, J. S. (1975). *The Relevance of Education.* Harmondsworth: Penguin, pp. xvi+205, £0·75.

BUROS, O. K. (Ed.) (1975). *Tests in Print* II. New Jersey : Gryphon Press, pp. 1,156, $70·00.

BURSTALL, C., JAMIESON, M., COHEN, S., and HARGREAVES, M. (1975). *Primary French in the Balance.* Windsor : NFER, pp. 304, n.p.

BUTTON, L. (1975). *Developmental Group Work with Adolescents.* London : Univ. of London Press, pp. xiv+208, p. £2·10, c. £3·95.

CANTWELL, Z. M., and HORTENSE, A. D. (1975). *Instructional Technology : An Annotated Bibliography.* New Jersey : Scarecrow Press, pp. vi+386, $12·50.

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (1975). *L'Annee Psychologique*, 1974. Paris : Presses Universitaires de France, pp. 691, n.p.

CLARKE, A. M., and CLARKE, A. D. B. (Eds.) (1975). *Mental Deficiency : The Changing Outlook.* London : Methuen, pp. xx+886, £16·50.

CHAZAN, M., MOORE, T., WILLIAMS, P., and WRIGHT, J. (1975). *The Practice of Educational Psychology.* London : Longman, pp. x+422, £5·99.

COAN, R. W. (1974). *The Optimal Personality: An Empirical and Theoretical Analysis.* London : Routledge, pp. x+242, £4·75.

COLLIER, G., WILSON, J., and TOMLINSON, P. (1975). *Values and Moral Development in Higher Education.* London : Croom Helm, pp. iii+225, £4·75.

CROCKER, A. C. (1975). *Predicting Teaching Success.* Windsor : NFER, pp. ix+215, n.p.

CROSS, G. R. (1974). *The Psychology of Learning : An Introduction for Students of Education.* Oxford : Pergamon Press, pp. viii+294, p. £3·80, c. £6·00.

FRASER, M. (1974). *Children in Conflict.* Harmondsworth : Penguin, pp. 208, £0·50.

FREUD, A. (1975). *Introduction to Psycho-Analysis : Lectures for Child Analysts and Teachers.* London : Hogarth Press, pp. xx+200, £3·75.

GETZ., W., WIESE, A. E., SUE, S., and AYERS, A. (1975).. *Fundamentals of Crisis Counselling.* London : Lexington, pp. xiv+184, £6·15.

GOLICK, M. (1975). *Learning Through Card Games.* London : Wolfe Publishing, pp. vi+100, £2·50.

GULLIFORD, R., and WIDLAKE, P. (1975). *Teaching Materials for Disadvantaged Children : Schools Council Bulletin 5.* London : Evans/Methuen, pp. 158, n.p.

Herriot, P. (1974). *Attributes of Memory.* Andover : Methuen, pp. x+205, p. £1·40, c. £3·25.

Holden, J., and Tyfield, L. (1975). *The Child with Phenylketonuria.* London : National Society for Mentally Handicapped Children, pp. 9, £0·15.

Inhelder, B., Sinclair, H., and Bovet, M. (1975). *Learning and the Development of Cognition.* London : Routledge, pp. xiv+308, £4·95.

Institut für Arbeitsmarkt und Berufsforschung (1974). *Literatur-Dokumentation.* Nurnberg, pp. v+1400.

Invalid Children's Aid Association (1974). *Say It in Pictures.* London : ICAA.

Jennings, M. K., and Niemi, R. G. (1974). *The Political Character of Adolescence : The Influence of Families and Schools.* New Jersey : Princeton Univ. Press, pp. 348, £6·60.

Joint Council for the Education of Handicapped Children (1975). *Integration or Segregation ? A False Alternative.* London : National Council for Special Education, pp. 4, £0·15.

Lambert, R. (1975). *The Chance of a Lifetime ?* London : Weidenfeld and Nicolson, pp. ix+423, £7·00.

Laufer, M. (1975). *Adolescent Disturbance and Breakdown.* Harmondsworth : Penguin, pp. xiii+96, £0·45.

Leach, P. (1974). *Babyhood.* Harmondsworth : Penguin, pp. 481, £0·90.

Lewis, D. G. (1975). *Assessment in Education.* London : Univ. of London Press, pp. ix+198, p. £1·90, c. £3·75.

Lewis, M., and Rosenblum, L. A. (Eds.) (1974). *The Effect of the Infant on its Caregiver.* Chichester : John Wiley, pp. xxiv+264, £8·00.

Lytton, H., and Craft, M. (Eds.) (1975). *Guidance and Counselling in British Schools.* London : Edward Arnold, pp. 189, £2·25.

Mann, L., and Sabatino, D. A. (Eds.) (1974). *The First Review of Special Education, Vols. I and II.* Pennsylvania : JSE Press, pp. iv+292, n.p.

Mann, L., and Sabatino, D. A. (Eds.) (1974). *The Second Review of Special Education.* Pennsylvania : JSE Press, pp. viii+405 n.p.

Marks, P. A., Seeman, W., and Haller, D. L. (1975). *The Actuarial Use of the MMPI with Adolescents and Adults.* Baltimore : Williams and Wilkins, pp. xix+324, $14·74.

Mehrabian, A., and Ksionzky, S. (1975). *A Theory of Affiliation.* Farnborough: Lexington, pp. xvi+212, £7·30.

Miles, T. R. (1975). *The Dyslexic Child.* London : Priory Press, pp. xiii+140, £3·50.

Miller, F. J. W., Court, S. D. M., Knox, E. G., and Brandon, S. (1974). *The School Years in Newcastle-upon-Tyne.* London : Oxford Univ. Press., pp. xvii+362, £6·00.

Mitchell, R. (1975). *Depression.* Harmondsworth : Penguin, pp. 112, £0·50.

Open University (1974). Milton Keynes. *Introduction to Planning and Decision Models*, pp. 45 £1·00 ; *Problems and Applications*, pp. 117, £1·80.

Oppenheim, A. N., and Torney, J. (1975). *The Measurement of Children's Civic Attitudes in Different Nations.* London : John Wiley, pp. xi+84, n.p.

Oxford Review of Education (1975). *Volume* 1, *Number* 1. Oxford : Carfax, pp. 89, £2·50.

Padgham, C. A., and Saunders, J. E. (1975). *The Perception of Light and Colour.* London : G. Bell, pp. viii+192, £6·50.

Peters, R. S. (1974). *Psychology and Ethical Development.* Hemel Hempstead : Allen and Unwin, pp. ix+480, £7·15.

Piaget, J., and Inhelder, B. (1974). *The Child's Construction of Quantities.* London : Routledge, pp. viii+285, £5·75.

Poteet, J. A. (1974). *Behaviour Modification : A Practical Guide for Teachers.* London : Univ. of London Press, pp. viii+104, £1·25.

Rhodes, W. C. (1974). *A Study of Child Variance, Vol.* 3 : *Service Delivery Systems.* Michigan : Univ. of Michigan, pp. xii+666, n.p.

Rowntree, D. (1975). *Educational Technology in Curriculum Development.* London : Harper and Row, pp. x+197, n.p.

Sandels, S. (1974). *The Skandia Report II : Why Are Children Injured in Traffic ? Can we Prevent Child Accidents in Traffic ?* Stockholm : Skandia, pp. 87, n.p.

370.1505/B777

Sauber, S. R. (1975). *Preventive Educational Intervention for Mental Health.* Chichester : John Wiley, pp. ixv+175, £4·80.

Schiefelbusch, R. L., and Lloyd, L. L. (Eds.) (1974). *Language Perspectives—Acquisition, Retardation and Intervention.* London : Macmillan, pp. xv+670, £7·95.

Schonell, F. J., and Goodacre, E. (1974). *The Psychology and Teaching of Reading.* Edinburgh : Oliver and Boyd, pp. 230, n.p.

Schools Council (1974). *Nuffield Secondary Science Evaluation.* London : Macmillan, pp. ix+70, £2·95.

Social Science Research Council (1975). *Inventory* 1974-75 : *SSRC Survey Archive.* Colchester : Univ. of Essex for SSRC, pp. 215, n.p.

Sprinthall, R. C., and Sprinthall, N. A. (1974). *Educational Psychology : A Developmental Approach.* London : Addison-Wesley, pp. viii+462, £6·05.

Stanley, J. C., Keating, D. P., and Fox, L. H. (Eds.) (1975). *Mathematical Talent: Discovery, Description and Development.* London : John Hopkins Univ. Press, pp. xvii+215, £5·50.

Szasz, T. (1974). *Ideology and Insanity.* Harmondsworth : Penguin, pp. 253, £0·60.

Thornburg, H. D. (1974). *Preadolescent Development.* Arizona : Arizona Univ. Press., pp. xii+356, $7·45.

Venables, E. (1975). *Intelligence and Motivation Among Day-Release Students.* Windsor : NFER, pp. xi+173, n.p.

Weeks, T. E. (1975). *The Slow Speech Development of a Bright Child.* Farnborough : Lexington, pp. xix+158, £6·85.

Winnicott, D. W. (1975). *Through Paediatrics to Psycho-Analysis.* London : Hogarth Press. ixl+350, £5·00.

ERRATUM

Wedell, K. (1973). *Learning and Perceptuo-Motor Disabilities in Children.* Chichester: Wiley, pp. 345, is priced at £2·75, not £8·00.

P8576

Problems of Language and Learning

Edited by **Alan Davies**

An authoritative symposium on the inter-relationship of language and learning. The Edinburgh seminar on which this book is based brought together leading international authorities and reviewed some significant current trends in sociolinguistics, psycholinguistics, and related fields. The contributors include M. A. K. Halliday on sociolinguistics ; J. S. Bruner on language as an instrument of thought ; Margaret Clark on the present state of research into reading problems ; James Britton on the teaching of writing ; and H. H. Speitel on dialect. There is a general summing-up by Basil Bernstein.
£3.00

The Intelligence of Feeling

Robert R. Witkin

"... provides an analysis in some depth into the consciousness and practice of the teachers of the creative arts in the light of Mr Witkin's own conceptual framework concerning the function of expressive activity in the developing experience of the child ... he has provided teachers with some disturbing data which will, with luck, make many re-examine their practice."
Times Educational Supplement
£3.75

Heinemann Educational Books
48 Charles Street,
London W1X 8AH

40 years of continuous service

for Discriminating Test Users

Edited by OSCAR K. BUROS

NEW Tests in Print II.

December 1974, 1146 pages; $70 A comprehensive index to tests, test reviews, and the literature on specific tests Presents 2,467 tests, 16,574 new references on the construction, use, and validity of specific tests, a cumulative name index for each test with references, the 1974 *Standards for Educational and Psychological Tests*, a directory of 493 publishers with complete listings of their tests, a title index which includes out of print tests, an author index to over 70,000 documents (tests, reviews, excerpts, and references), and a scanning index for locating tests for a particular population.

1972 The Seventh Mental Measurements Yearbook

1972; 2032 pages in 2 volumes, $70 per set. "Indispensable . comprises all the values of its predecessors, but in larger measure . simple, clear indexing system .. statistics can give only the most inadequate picture the newest edition lists 1,157 tests with 12,372 references For 546 of the tests, there are 798 original reviews by 439 reviewers, supplemented by another 181 reviews excerpted from 39 journals" — *Test Service Bulletin of The Psychological Corporation,* May 1972.

1970 Personality Tests and Reviews

1970, 1690 pages, $45 00. "The second of four monographs covering specific sections of the first six *Mental Measurements Yearbooks.* * A great deal of new material is also presented for the first time." — Wayne H Holtzman, *Am J Orthopsychiatry,* July 1970

1968 Reading Tests and Reviews

1968, 542 pages; $20.00. "Brings together in convenient form the enormous amount of bibliographical material about reading tests of the past and present that have been available heretofore only by consulting separately the six *Mental Measurements Yearbooks.*" — Frederick B. Davis, *Ed & Psychol Meas,* Autumn 1969.

Earlier Volumes Needed for Full Coverage

1965 **The Sixth Mental Measurements Yearbook.** Reprinted 1971, 1751 pages, $45 00
1961 **Tests in Print.** Reprinted 1974, 508 pages, $15 00
1959 **The Fifth Mental Measurements Yearbook.** 1321 pages, $35.00
1953 **The Fourth Mental Measurements Yearbook** Reprinted 1974, 1188 pages, $30 00
1949 **The Third Mental Measurements Yearbook.** Reprinted 1974, 1062 pages, $25 00
1941 **The 1940 Mental Measurements Yearbook.** Reissued 1972, 672 pages, $20 00
1938 **The 1938 Mental Measurements Yearbook.** Reissued 1972, 430 pages, $17 50

No postage charges on prepaid orders

Order directly from
THE GRYPHON PRESS 220 Montgomery Street, Highland Park, N. J. 08904

RECENTLY PUBLISHED

Behavior Modification

A Practical Guide for Teachers

James A. Poteet, Ph.D., Director of the Diagnostic Teaching Centre, Indianapolis.

This readable and well-organised book provides a sound, practical introduction to the techniques and uses of behaviour modification with minimal reference to theory. Set in the US school context, it covers issues of general importance to all teachers and will stimulate analysis of teaching methods and techniques. With the goal of a ' positive classroom environment,' the reader learns to select and define desired behaviour and then to measure, evaluate and modify existing behaviour. In addition to developing these skills, he covers such topics as stimulus control, attitude change, exceptional children and accountability.

Unibook £1.25 **0 340 18499 X**

Lecturers are invited to write for inspection copies on approval.

UNIVERSITY OF LONDON PRESS LTD

Dept P971, St. Paul's House,
Warwick Lane, London EC4P
4AH.

PSYCHOLOGY IN GREAT BRITAIN

Edited by B. M. FOSS

Issued as a special Supplement to the *Bulletin* of the British Psychological Society on the occasion of the XIXth International Congress of Psychology, *Psychology in Great Britain* provides a useful survey of developments in British psychology and related fields.

Obtainable, price 25p post free, from
THE BRITISH PSYCHOLOGICAL SOCIETY
18-19 Albemarle Street, London W1X 4DN

PSYCHOLOGY IN THE FUTURE

A Symposium reprinted from the Proceedings of the XIX International Congress of Psychology, London, 1969.

Contents

Preface

Chairman's Introduction. By George A. Miller

Psychology Tomorrow—The Immediate Future
By Karl H. Pribram

Psychology A.D. 2000 : Facts, Forecasting, Fantasies and Fallacies
By Harry Kay

Possible Roles of Psychology in the Very Distant Future
By Masanao Toda

Obtainable, price 15p post free, from
THE BRITISH PSYCHOLOGICAL SOCIETY
18-19 Albemarle Street, London W1X 4DN

INTERNATIONAL JOURNAL OF PSYCHOLOGY
JOURNAL INTERNATIONAL DE PSYCHOLOGIE

Published quarterly by the International Union of Psychological Science

The International Journal of Psychology is devoted to cross-cultural comparisons of psychological phenomena in general psychology, child psychology and social psychology. It also deals with international confrontation of theories and methods on which psychologists from various countries differ because of their native culture and training. Oriented towards basic research and theory, it publishes scientific papers in English or in French. In each issue, an International Platform gives news and informations on internationally significant problems.

Editor (Scientific papers)	J. Leroux 100, Tiensestraat 3000 Louvain, BELGIUM
Associate Editor (International Platform)	W. Holtzman University of Texas, Austin, U.S.A.

1974 VOLUME 9 No. 1

CONTENTS **SOMMAIRE**

The questionable unity of the concrete operations stage
ALASTAIR HERON AND WENDY DOWEL 1

Social conformity and psychological differentiation
HERMAN A. WITKIN, DOUGLASS PRICE-WILLIAMS, MARIO BERTINI, BJORN CHRISTIANSEN, PHILIP K. OLTMAN, MANUEL RAMIREZ and JACQUES VAN MEEL 11

The influence of culture and sex on the perception of persons
FRANCES F. KORTEN 31

Internal versus external control in India and Canada
DAVID W. CARMENT 45

A note on the cross-national differences in cognitive conflict found by Hammond et al.
BERNDT BREHMER 51

The semantics of asking a favour. An attempt to replicate cross-culturally
JOHN M. INNES 57

The 1974 Directory of Cross-Cultural Research and Researchers 30

INTERNATIONAL PLATFORM FOR PSYCHOLOGISTS — TRIBUNE INTERNATIONALE DES PSYCHOLOGUES

Directory of IUPS Member Societies 65

Annual subscription (4 issues)	**$16.00**
Single issue	**4.00**

Should be sent to : **Centrale des Revues DUNOD-GAUTHIER-VILLARS**
24-26, Boulevard de l'Hôpital 75005 PARIS

EDUCATIONAL RESEARCH

Vol. 17, No. 2 *Editor* : ALFRED YATES *February*, 1975

C. M. FERGUSON and S. GARRETT. Stimulus material in 'O' Level History.

J. NISBET. Adding and averaging grades.

P. L. GARDNER. Attitude measurement : a critique of some recent research.

P. O. OKUNROTIFA. Attitudes of Nigerian secondary school children to programmed instruction in Geography.

L. B. HENDRY. The role of the physical education teacher.

E. CARRANZA. The impact of teacher life changes and performance on student dropouts.

B. THOMPSON. Adjustment to school.

B. A. AKHURST. The prevalence of behaviour problems among children in care.

G. FENWICK. Junior school pupils' rejection of school library books.

BOOK REVIEWS.

The subscription is £2·75 a year (incl. postage); £3·20 Overseas (postage paid).

All orders should be sent to :

NFER Publishing Company Ltd. (Journals Department), 2 Jennings Buildings,
Thames Avenue, Windsor, Berks., SL4 1QS

The British Journal of Medical Psychology

Volume 48, Part 1, March 1975

Eleanor S. Wertheim. Person-environment interaction : the epigenesis of autonomy and competence. I. Theoretical considerations (normal development)
Nils Cochrane. Assessing the aggressive component of personality. I. Problems
Nils Cochrane. Assessing the aggressive component of personality. II. A new approach
Alexander P. W. Shubsachs. To repeat or not to repeat ? Are frequently used constructs more important to the subject ? A study of the effect of allowing repetition of constructs in a modified Kelly Repertory Test
Anthony Ryle and **Susan Lipshitz.** Recording change in marital therapy with the reconstruction grid
David G. McDonald and **Elizabeth A. King.** Measures of sleep disturbance in psychiatric patients
Marcia D. Mather and **Gian S. Degun.** A comparative study of hypnosis and relaxation
Lester Luborsky, Loren Crabtree, Homer Curtis, George Ruff and **Jim Mintz.** The concept 'space' of transference for eight psychoanalysts
Benjamin Beit-Hallahmi and **Michael Argyle.** God as a father-projection : the theory and the evidence
Anthony Fry. Explanation and understanding in general psychiatry : two fundamental aspects of the treatment and approach to the patient
Books Received and Short Book Notices

£4.00 net (US $12.00 in USA and Canada). 1975 subscription £13.00 net (US $39.00 in USA and Canada)

CAMBRIDGE UNIVERSITY PRESS
BENTLEY HOUSE, 200 EUSTON ROAD, LONDON NW1 2DB
32 EAST 57TH STREET, NEW YORK, N.Y. 10022

The British Journal of Psychology

Volume 66 Part I February, 1975

SIMON FOLKARD. Diurnal variation in logical reasoning

A. B. HILL. Extraversion and variety seeking in a monotonous task

J. F. ALLSOPP and H. J. EYSENCK. Extraversion, neuroticism and verbal reasoning ability as determinants of paired-associates learning

CHARLES A. PADGHAM. Colours experienced in dreams

HADYN ELLIS, JOHN SHEPHERD, and GRAHAM DAVIES. An investigation of the use of the Photo-fit technique for recalling faces

KENNETH FLOWERS. Handedness and controlled movement

ELIZABETH R. L. McMEEKAN and W. A. LISHMAN. Retest reliabilities and interrelationship of the Annet Hand Preference Questionnaire and the Edinburgh Handedness Inventory

D. R. DAVIES and D. M. JONES. The effects of noise and incentives upon attention in short-term memory

GRAHAM QUINTON and BRIAN J. FELLOWS. 'Perceptual' strategies in the solving of three-term series problems

J. G. WOLFF. An algorithm for the segmentation of an artificial language analogue

LUBOMIR S. PRYTULAK and DELBERT A. BRODIE. Effect of length, density and angle between arms on gestalt grouping

EUGENE C. LECHELT Temporal numerosity discrimination : intermodal comparisons revisited

BOOK REVIEWS

OTHER PUBLICATIONS RECEIVED

£5·00 net ($15·00 in USA and Canada). 1975 subscription £15·00 net ($45·00 in USA and Canada).

CAMBRIDGE UNIVERSITY PRESS

Bentley House, 200 Euston Road, London, NW1 2DB
32 East 57th Street, New York, N.Y. 10022

THE BRITISH JOURNAL OF SOCIAL AND CLINICAL PSYCHOLOGY

Volume 14 Part I February, 1975

KEITH GIBBINS and TONYA K. GWYNN. A new theory of fashion change : a test of some predictions

ANGELA B. SUMMERFIELD. Errors in decoding tone of voice during dyadic interaction

MARK COOK and JACQUELINE M. C. SMITH. The role of gaze in impression formation

JUDY TODD. Social evaluation orientation, task orientation and deliberate cueing in experimenter expectancy effect

LUC M. LEFEBVRE. Encoding and decoding of ingratiation in modes of smiling and gaze

W. LAUTERBACH. Assessing psychological conflict

W. LAUTERBACH. Covariation of conflict and mood in depression

PETER W. BRADSHAW, PHILIP LEY, JOHN A. KINCEY and JENNIFER BRADSHAW. Recall of medical advice : comprehensibility and specificity

ALFRED B. HEILBRUN, JR. A proposed basis for delusion formation within an information—processing model of paranoid development

EDGAR MILLER. Impaired recall and the memory disturbance in presenile dementia

D. R. HEMSLEY. A two-stage model of attention in schizophrenia research

LARRY C. JENSEN and ANN M. RYTTING. Changing children's beliefs about punishment

PETER G. COLEMAN. Interest in personal activities and degree of perceived implications between personal constructs

D. R. THOMAS. Conservatism, authoritarianism and child-rearing practices

J. P. N. PHILLIPS. A note on the scoring of the grid test of schizophrenic thought disorder

BOOK [illegible]

OTHER PUBLICATIONS RECEIVED

£4·50 net ($13·50 in USA & Canada). 1975 subscription £14·00 net ($42·00 in USA & Canada)

CAMBRIDGE UNIVERSITY PRESS

BENTLEY HOUSE, 200 EUSTON ROAD, LONDON, NW1 2DB
32 EAST 57th STREET, NEW YORK, N.Y. 10022

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology
The British Journal of Medical Psychology
The British Journal of Social and Clinical Psychology
The British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology (Jointly with A.T.C.D.E.)

The British Journal of Psychology, The British Journal of Medical Psychology and *The British Journal of Social and Clinical Psychology* appear in quarterly parts, the 1975 subscription price, payable in advance, being £15·00 (US $45·00 in the USA and Canada) for *The British Journal of Psychology*, £14·00 (US $42·00) for *The British Journal of Social and Clinical Psychology* and £13·00 (US $39·00) for *The British Journal of Medical Psychology*. Subscriptions to any of these journals may be sent to any bookseller or subscription agent, or to the Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB. American and Canadian enquiries should be addressed to the Cambridge University Press, American Branch, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November. The subscription rate, payable in advance, is £10·00 per volume post free (US $30·00 in the USA). Subscriptions should be sent to the Society's agents, John Wright and Sons, Ltd., 42-44 Triangle West, Bristol, BS8 1EX, or to any bookseller.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The prepaid subscription is £4 a year post free, or £1·50 each number. The prepaid subscription price in the USA is $13·50 post free or $4·50 each number. Subscriptions should be sent to the Scottish Academic Press, 25 Perth Street, Edinburgh EH3 5DW.

The series of *Monograph Supplements* issued in connection with *The British Journal of Psychology* is now discontinued. Orders may be sent to any bookseller, or to Cambridge University Press, Bentley House, 200 Euston Road, London, NW1, from whom details of these publications may be obtained.

MANUSCRIPTS

For *The British Journal of Psychology* should be sent to Professor A. D. B. Clarke, Department of Psychology, The University, Hull, HU6 7RX ;

For *The British Journal of Medical Psychology* should be sent to Professor Arthur H. Crisp, Psychiatric Research Unit, Atkinson Morley's Hospital, Copse Hill, London, SW20 0NE ;

For *The British Journal of Social and Clinical Psychology* should be sent to Dr. Halla Beloff (Social Psychology Editor), Department of Psychology, University of Edinburgh, 60 Pleasance, Edinburgh, EH8 9TJ ; or to Dr. H. R. Beech (Clinical Psychology Editor), Department of Psychiatry, Withington Hospital, West Didsbury, Manchester, M20 8LR ;

For *The British Journal of Mathematical and Statistical Psychology* should be sent to Professor P. Levy, Department of Psychology, University of Lancaster, Lancaster, LA1 4YF ;

For *The British Journal of Educational Psychology* should be sent to Professor N. J. Entwistle, Department of Educational Research, University of Lancaster, Lancaster.

Contributors are advised to consult *Suggestions to Authors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price 25p (USA $1·00) post free, from Cambridge University Press.

OFF-PRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £12·00 (£10·00 if paid by Banker's Order).

They are entitled to purchase any or all of the following Journals at the special rates of £1·50 per Journal per annum : *The British Journal of Psychology, The British Journal of Medical Psychology, The British Journal of Social and Clinical Psychology, The British Journal of Mathematical and Statistical Psychology* and *The British Journal of Educational Psychology*.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, 18-19 Albemarle Street, London, W1X 4DN.

THE ASSOCIATION OF TEACHERS IN COLLEGES AND DEPARTMENTS OF EDUCATION

Education for Teaching is published three times a year at mid-term in November, February and May, and distributed free to members of the Association. To non-members it costs 25p per issue, post free, (subscription 70p per annum).

Members of the A.T.C.D.E. receive *The British Journal of Educational Psychology* at a reduced rate.

Information on membership and further activities may be obtained from the Honorary Secretary, 151 Gower Street, London, WC1.

17 SEP 1975

CONTENTS

KALLOS, D., and LUNDGREN, U. P.: Educational Psychology: Its Scope and Limits . . . 111

BOYDELL, D.: Pupil Behaviour in Junior Classrooms 122

TROWN, E. A., and LEITH, G. O. M.: Decision Rules for Teaching Strategies in Primary Schools: Personality-Treatment Interactions 130

ROBINSON, W. P.: Boredom at School 141

STOTT, D. H., and ALBIN, J. B.: Confirmation of a General Factor of Effectiveness-Motivation by Individual Tests 153

BROADBENT, D. E.: Cognitive Psychology and Education 162

PEEL, E. A.: Predilection for Generalising and Abstracting. . . 177

KELLY, M., and PHILP, H.: Vernacular Test Instructions in Relation to Cognitive Task Behaviour Among Highland Children of Papua New Guinea 189

VALENTINE, E. R.: Performance on two Reasoning Tasks in Relation to Intelligence, Divergence and Interference Proneness . 198

CHOPPIN, B.: Guessing the Answer on Objective Tests 206

MUKHERJEE, B. N.: The Factorial Structure of Wechsler's Pre-School and Primary Scale of Intelligence at Successive Age Levels . 214

RESEARCH NOTES:

BANRETI-FUCHS, K. M.: Attitudinal, Situational and Mental Health Correlates of Academic Achievement at the Undergraduate University Level 227

ROWELL, J. A., and RENNER, V. J.: Personality, Mode of Assessment and Student Achievement 232

BOOK REVIEWS 237

PUBLICATIONS RECEIVED 253

UK ISSN 0007–0998